CIS 624: Structure of Programming Languages

Lecture 1 — Course Introduction

Boyana Norris
2017
Today

- Administrative stuff
- Course motivation and goals
  - A Java example
- Course overview
- Course pitfalls
- Start OCaml tutorial (see separate notes)
  - Advice: start playing with it soon (e.g., hw1, problem 1)
Course facts

- Instructor: Boyana Norris, norris@cs.uoregon.edu
- Office hours: 307 Deschutes (see course web page for times), email for appointments at other times (any other time – if my door is open, please drop in!)
- Course web page:
  https://www.cs.uoregon.edu/Classes/17F/cis624
Coursework

- 3–5 homework assignments
  - “Paper/pencil” (\LaTeX recommended?)
  - Programming (OCaml required)
  - Where you’ll probably learn the most
  - Do challenge problems if you *want* but not technically “extra”
  - First homework carefully pipelined with lectures

- One “introduction/summary” to a published research paper
  - More details later; high work/length ratio

- Two exams
  - Samples will be provided
Academic integrity

▶ Don’t cheat in my class
  ▶ I’ll be personally offended
  ▶ Being honest is far more important than your grade

▶ Rough guidelines
  ▶ Can sketch idea together
  ▶ Cannot look at code solutions

▶ Ask questions and always describe what you did

▶ Please *do* work together and learn from each other
Logistical Advice

- Reading
  - Slides/proofs posted, but they are enough to teach from not to learn from (note that the references will be updated during the first couple of weeks)
  - Consider getting the book(s), but request additional information as needed
  - Will work through examples by hand (mine, sometimes yours)

- Work together (while maintaining academic integrity)
- For some assignments, you will be directed to work in teams of two to three people. You can still choose to work alone if you prefer.

- Arrive on time
- Ask questions
Programming-language concepts

Focus on *semantic* concepts:

What do programs mean (do/compute/produce/represent)?

How to define a language *precisely*?

English is a poor *metalanguage*

**Definition**

A *metalanguage* is a form of language or set of terms used for the description or analysis of another language.

Other meta examples?
Aspects of meaning for PLs:

- equivalence, termination, determinism, type, . . .

This course does *not* give superficial exposure to *N* weird PLs
- But it will help you learn new languages via foundations
- And build rigorous models for any area of CS research

Boyana Norris

CIS 624 2017, Lecture 1

7
Does it matter?

Novices write programs that “work as expected,” so why be rigorous/precise/pedantic?

- The world runs on software
  - Web servers and nuclear reactors don’t “seem to work”

- You buy language implementations—what do they do?

- Software is buggy—semantics assigns blame

- Real languages have many features: building them from well-understood foundations is good engineering

- Never say “nobody would write that” (surprising interactions)

- ... in recent news
Is this Really about PL?

Building a precise model is a hallmark of quality research.

The value of a model is in its:
- Fidelity
- Convenience for establishing (proving) properties
- Revealing alternatives and design decisions
- Ability to communicate ideas concisely

Why we mostly do it for programming languages:
- Elegant things we all use
- Remarkably complicated (need rigor)

I believe this “theory” makes you a better computer scientist
- Focus on the model-building, not just the PL features
APIs

Like almost anything in computing, we can describe the course in terms of designing an API.

Many APIs have 1000s of functions with simple inputs:
- Kernel calls take a struct or two and return an int.

A typical language implementation more or less has just:
- `typecheck : program → bool`
- `compile : program → (string → value)`

But defining `program` and these functions is subtle, hard:
- Conversely, “a data structure is just a really dumb PL”
- Every extensible system ends up defining a PL (game engines, editors, web browsers, CAD tools, ...).
Java example

```java
class A { int f() { return 0; } }

class B {
    int g(A x) {
        try { return x.f(); }
        finally { s }
    }
}
```

For all $s$, is it equivalent for $g$’s body to be “return 0;”?
Motivation: code optimizer, code maintainer, ...
Punch-line

Not equivalent:

- Extend A
- \( x \) could be \texttt{null}
- \( s \) could modify global state, \textit{diverge}, throw, ...
- \( s \) could return

A silly example, but:

- PL makes you a good adversary, programmer
- PL gives you the tools to argue equivalence (hard!)
Course goals

1. Learn intellectual tools for describing program behavior

2. Investigate concepts essential to most languages
   - mutation and iteration
   - scope and functions
   - types
   - objects
   - threads

3. Write programs to “connect theory with the code”

4. Sketch applicability to “real” languages

5. Provide background for current PL research
   (less important for most of you)
Course nongoals (what we will not do)

- Study syntax; learn to specify grammars, parsers
  - Transforming $3 + 4$ or $(+ 3 4)$ or $+(3, 4)$ to “application of plus operator to constants three and four”
- Learn specific programming languages (but some OCaml)
What we will do

- Define really small languages
  - Usually Turing complete
  - Always unsuitable for real programming

- Extend them to realistic languages less rigorously

- Study models very rigorously via *operational models*

- Do programming assignments in OCaml
OCaml

- OCaml is an awesome (a.k.a. very cool), high-level language
- We will use a tiny core subset of it that is well suited for manipulating recursive data structures (like programs!)
- You mostly have to learn it outside of class
  - Don’t procrastinate
  - Don’t hesitate to ask questions
- Resources on course webpage
Pitfalls

How to hate this course and get the wrong idea:

▶ Forget that we made simple models to focus on the essence

▶ Don’t quite get inductive definitions and proofs when introduced

▶ Don’t try other ways to model/prove the idea
  ▶ You’ll probably be wrong
  ▶ And therefore you’ll learn more

▶ Think PL people focus on only obvious facts
  ▶ Need to start there
Final Metacomment

Acknowledging others is crucial...

This course is based on materials developed by Prof. Dan Grossman at the University of Washington and draws heavily on pedagogic ideas from at least: Chambers, Chong, Felleisen, Flatt, Fluet, Harper, Morrisett, Myers, Pierce, Rugina, Walker

And material covered in texts from Pierce, Wynskel, and others

(This is a course, not my work.)
"Let go of Java/C"

If you have seen SML, Haskell, Scheme, Lisp, etc. this will feel more familiar

If you have seen Caml, focus on the subset that will be most useful to us in studying PL

Give us some small code snippets so we have a common experience we can talk about