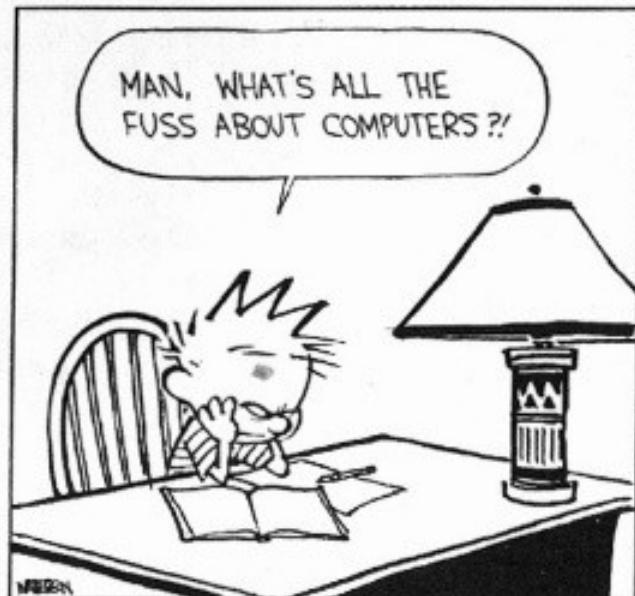
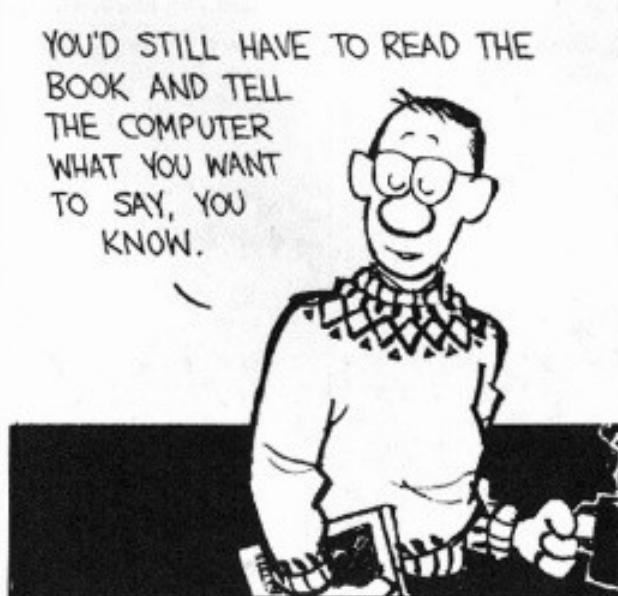
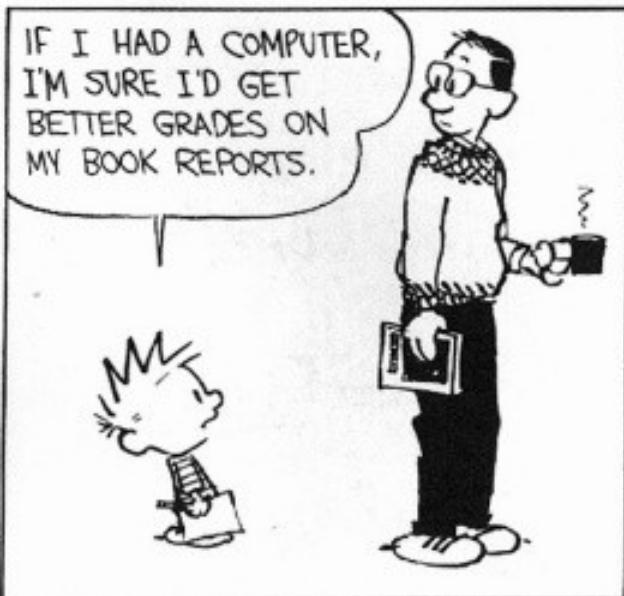


# Principles of Programming Languages

Wes Weimer

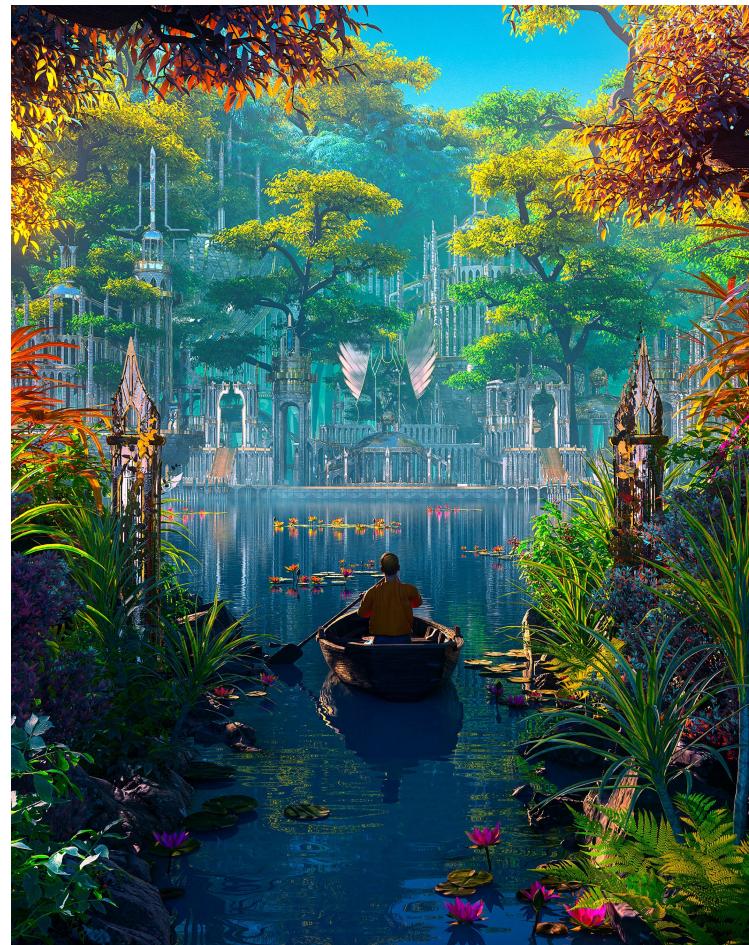
TR 2:50 - 4:15, Searles 223

w.weimer@bowdoin.edu



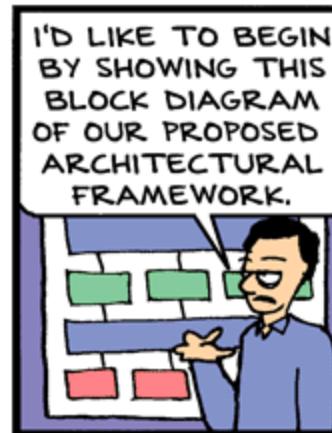
# Honestly

- It's going to be an awesome exploration ...



# Cunning Plan

- Administrivia
  - Webpage, Wes
- What Is This Class About?
- Easy or Hard? Work and Grading.
- Understanding a Program in Stages



# Course Home Page

- Find via Canvas or Piazza
- <https://weimer.github.io/csci2320/index.html>
- Lectures slides are available before class
  - You should still take notes! (citation next)
- Assignments are listed
  - also grading breakdown, regrade policies, etc.
- We make heavy use of the class forum

# Please Ask Me To Support Claims!

“...students who took notes on laptops performed worse on conceptual questions than students who took notes longhand. We show that whereas taking more notes can be beneficial, laptop note takers’ tendency to transcribe lectures verbatim rather than processing information and reframing it in their own words is detrimental to learning.”

[ Pam Mueller, Daniel Oppenheimer. *The pen is mightier than the keyboard: advantages of longhand over laptop note taking.* Psychol. Sci. 2014 Jun; 25(6):Epub 2014 Apr 23. ]

# Programming Languages

## Course Goals

- At the end of this course, you will be acquainted with the fundamental concepts in the **design** and **implementation** of high-level programming **languages**. In particular, you will understand the **theory** and **practice** of **lexing**, **parsing**, **semantic analysis**, and **code interpretation**. You will also have gained practical experience programming in **different languages**.

# What about you?

- What do you want to do or learn in this class?

The image shows a Twitter thread. The first tweet is from Mark Dalgleish (@markdalgleish) at 7:45 AM on 3/18/21 via Twitter for Android. It contains three messages:

- Young person: I think I might be interested in programming.
- University: Polymorphism.
- Young person: Maybe not.

The second tweet is a reply from Mark Dalgleish (@markdalgleish) 1 hour ago, replying to himself. It contains two messages:

- Internet tutorial: Here's how you make the computer go beep boop.
- Young person: Now we're talking.

At the bottom of the screenshot, there are engagement metrics: 2 replies, 4 retweets, 94 likes, and a share icon.

# Who Cares?

- In most cases, there is a clear mapping between **CS classes and jobs** or internships:
  - Take Databases → work at Oracle
  - Take OS → work at Microsoft
  - Take Financial ML → work in FinTech
  - Take PL → ???
- *Which companies develop compilers or interpreters?*

# Microsoft

- Visual Studio, Excel, etc.

The screenshot shows the Microsoft Visual Studio website with a purple header bar. The header includes links for Products, Features, Downloads, News (which is underlined), Support, Marketplace, Documentation, MSDN Subscriptions, Sign in, and a search icon. A green button on the right says "Free Visual Studio" with a circular arrow icon. The main content area has a purple background and features the heading "Visual Studio Team Services Jobs". Below it is the sub-headline "Become a member of the Visual Studio Team Services Team". A bulleted list of benefits follows: "✓ Work with the latest technologies", "✓ Work on a fast-paced agile team", and "✓ Have a big impact on the software industry through building an innovative service". At the bottom right of the purple area is a green button labeled "Send Resume >". Below the purple area, there are two small buttons: "Open Positions >" and "Life at Microsoft >".

Join us!

Become a key member of the Visual Studio Team Services (VSTS) service team led by [Brian Harry](#) and build the next generation of development services in the cloud! VSTS provides software development teams with version control, build automation, agile work management, social experiences and more to nearly 3,000,000 users.

# Oracle

- Java Compiler, Java Virtual Machine

The screenshot shows the Oracle Java Software homepage. At the top, there's a navigation bar with links for Sign In/Register, Help, Country, Communities, I am a..., I want to..., a search bar, and menu items for Products, Solutions, Support, Training, Partners, About, and OTN. A red banner on the left says "Java Software" and features the text "Create the Future". Below it, it says "Java is the world's #1 programming language". There are two buttons: "Java for Developers" with a download icon and "Java for Consumers" with a dropdown arrow. A video thumbnail for "Java Embedded for IoT (2:51)" is also present. On the right, there's a large image of a man looking at a futuristic city skyline with a "20 YEARS 1995-2015" banner and a Java logo. A sidebar on the left lists various roles: Java Developer, Database Administrator / Developer, System Admin / Developer, Architect, C-Level Executive (Chief Financial Officer, Chief Human Resources Officer, Chief Information Officer), and Other Roles (Analyst, Investor, Job Seeker, Partner, Student, Midsize Company). The "Job Seeker" link is underlined and has a cursor icon over it. At the bottom, there are tabs for Overview, Roles, Technologies, and Get Started.

# Intel

## • ICC

The screenshot shows the Intel Developer Zone homepage. At the top, there's a navigation bar with the Intel logo, "Developer Zone", "Join Today >", and "Log in". Below the navigation, there are links for "Development > Tools > Resources >". A search bar is on the right with the text "powered by Google" and a magnifying glass icon. The main content area has a dark header with the text "Intel® C++ Compilers".

### Leadership Application Performance

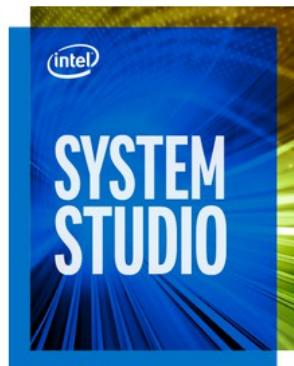
- Boost C++ application performance
- Future-proof code by making code that scales
- Plugs right into your development environment

If you are here, you are looking for ways to make your application run faster. Boost performance by augmenting your development process with the Intel® C++ Compiler. The Intel C++ Compiler plugs right into popular development environments like Visual Studio®, Eclipse®, XCode®, and Android Studio®. The Intel C++ Compiler is compatible with popular compilers including Visual C++® (Windows\*) and GCC (Linux®, OS X® and Android®).

The Intel C++ Compiler is available in four products based on your application development needs:



Intel® C++ Compiler in Intel® Parallel Studio XE



Intel® C++ Compiler in Intel® System Studio



Intel® C++ Compilers in Intel® INDE (support only)



Intel® Bi-Endian Compiler

# Google

- Go, Dart, etc.

The screenshot shows two side-by-side web pages. On the left is the Dart website, featuring a blue header with the Dart logo and navigation links for 'GET STARTED', 'FUNDAMENTALS', 'WEB', 'SERVER', and 'MORE'. A dropdown menu from the 'MORE' link is open, showing options like 'Code Samples', 'Synonyms with Other Languages', 'Dart by Example', 'Articles', 'Language Specification', 'Who Uses Dart', 'FAQ', and 'Logos and Colors'. Below the header, the text 'Scalable, productive' is displayed, followed by a description of Dart as an open-source language. A 'TRY DART' button is visible. On the right is the Go Programming Language landing page, with a header that includes the Dart logo and navigation links for 'Documents', 'Packages', 'The Project', 'Help', 'Blog', and 'Search'. Below the header, there's a section titled 'Try Go' containing a code editor with Go code for printing 'Hello, 世界'. At the bottom of this section are buttons for 'Hello, World!', 'Run', 'Share', and 'Tour'. To the right of the code editor is a large text block about Go and a cartoon gopher character. A 'Download Go' button is at the bottom right, with a note about binary distributions for Linux, Mac OS X, Windows, and more.

Dart

GET STARTED ▾ FUNDAMENTALS ▾ WEB ▾ SERVER ▾ MORE ▾

Scalable, productive

Dart is an open-source, scalable programming language, with robust tools for building web, mobile, server, and mobile apps.

TRY DART

Code Samples  
Synonyms with Other Languages  
Dart by Example  
Articles  
Language Specification  
Who Uses Dart  
FAQ  
Logos and Colors

The Go Programming Language

Documents Packages The Project Help Blog Search

Try Go

// You can edit this code!  
// Click here and start typing.  
package main

```
import "fmt"

func main() {
    fmt.Println("Hello, 世界")
}
```

Pop-out ↗

Hello, World! Run Share Tour

Go is an open source programming language that makes it easy to build simple, reliable, and efficient software.

Download Go

Binary distributions available for Linux, Mac OS X, Windows, and more.

# Wind River, Green Hills

## • Embedded!

### DIAB COMPILER

For over 25 years, Wind River Diab Compiler has been helping industrial, medical, and aerospace industries. Diab Compiler generates extremely fast, high-quality object code in the smallest possible footprint, and produce high-quality, standards-compliant code.



Diab Compiler's unique optimization technology generates extremely fast, high-quality object code in the smallest possible footprint.



The Latest Inc

Due to collaboration with our partners, Diab Compiler is now available for many microcontrollers and older processors. Due to time to allow for the development of the compiler for these target systems, we are currently hiring for the following groups:



**Leading the Embedded World**



#### Jobs - Opportunities in the USA

Green Hills Software is always looking for qualified Engineering, Sales, and Marketing staff. Please submit your resume to the Corporate Office where it will be processed and reviewed by the hiring manager.

Click on a job title below for a complete description of the position:

- Corporate Field Applications Engineer (Santa Barbara, CA)
- Embedded Software Consultant (Santa Barbara, CA)
- Embedded Solutions Test Engineer (Santa Barbara, CA)
- Field Engineer (Santa Barbara, CA)
- Field Services Engineer (Santa Barbara, CA)
- Functional Safety Software Engineer (Santa Barbara, CA)
- Product Engineer (Santa Barbara, CA)
- Sales Managers (location TBD)
- Software Development Engineer (Santa Barbara, CA)
- Technical Marketing Engineer (Santa Barbara, CA)

[Click here for information on applying.](#)

Green Hills Software is an Equal Opportunity / Affirmative Action Employer.



**Software Development Engineer (Santa Barbara, CA)**

**Job description:**

A software engineer has complete engineering responsibility for one or more major components of the Green Hills product line. For an experienced programmer this is a satisfying position in which you have personal responsibility for creating a tool used by thousands of programmers around the world. Our engineers are involved in Language Front Ends, Code Generators, Real Time Operating Systems, our MULTI Development Environment, our Secure Workstation, and Target Systems.

Here are the groups for which we are hiring:

- **Compiler:** Create, update, and maintain a language front end or a target architecture backend for the highly-optimizing family of Green Hills compilers. A compiler engineer might work on new language extensions, specific cutting-edge optimizations for the latest chips to hit the market, or on general optimizations that will benefit our entire product line. An ideal candidate understands low level microarchitecture designs and is comfortable working with assembly code, yet can also develop tools written in high level languages.

# Wait, what? Embedded?

- Curiosity Mars Rover, Cell Phones, Satellites, Engine Control Modules, Computed Radiology, Fighter Jets, Digital Cameras, Turbines, Anti-Lock Brakes, Wii U Game Console, Nintendo Switch, ...



Green Hills Software's MULTI Integrated Development Environment Selected by Nintendo for Wii U Development

Global Agreement Will Yield Richer Games, with Faster Time-to-Market

SAN JOSE, CA — March 27, 2012 — DESIGN West/ESC 2012, Booth #1227 — Green Hills Software, the largest independent vendor of embedded software solutions, has entered into a global license agreement that will enable Nintendo Co., Ltd. to provide Green Hills Software's MULTI® integrated development environment (IDE) to developers that are creating video game software for the upcoming Wii U platform, which is scheduled to be launched later this year.

"We selected the Green Hills Software solution because it generates highly optimized code, and Green Hills provides excellent global support," commented Mr. Genyo Takeda, senior managing director of Integrated Research & Design at Nintendo Co., Ltd.



of supporting life and to assess the planet's habitability for future human missions.

# Adobe

- Photoshop contains interpreters ...

## 2 Photoshop Scripting Basics

This chapter provides an overview of scripting for Photoshop, describes scripting support for the scripting languages AppleScript, VBScript, and JavaScript, how to execute scripts, and covers the Photoshop object model. It provides a simple example of how to write your first Photoshop script.

If you are familiar with scripting or programming languages, you most likely will want to skip much of this chapter. Use the following list to locate information that is most relevant to you.

- For more information on the Photoshop object model, see ["Photoshop Object Model" on page 11](#).
- For information on selecting a scripting language, refer to the *Introduction to Scripting* guide.
- For examples of scripts created specifically for use with Photoshop, see Chapter 3, ["Scripting Photoshop" on page 21](#).
- For detailed information on Photoshop objects and commands, please use the reference information in the three reference manuals provided with this installation: *Adobe Photoshop CC 2015 AppleScript Scripting Reference*, *Adobe Photoshop CC 2015 Visual Basic Scripting Reference*, and *Adobe Photoshop CC 2015 JavaScript Scripting Reference*.

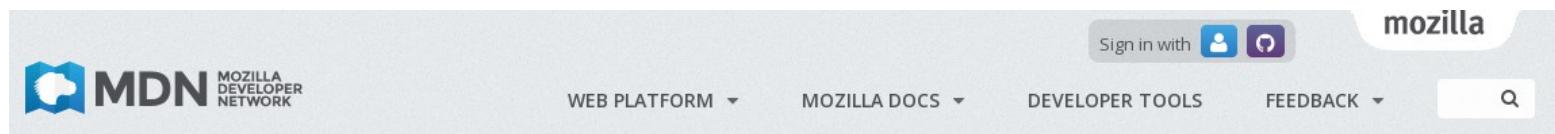
**NOTE:** You can also view information about the Photoshop objects and commands through the object browsers for each of the three scripting languages. See ["Viewing Photoshop Objects, Commands, and Methods" on page 21](#).

### Scripting Overview

A script is a series of commands that tells Photoshop to perform a set of specified actions, such as applying different filters to selections in an open document. These actions can be simple and affect only a single object, or they can be complex and affect many objects in a Photoshop document. The actions can call Photoshop alone or invoke other applications.

# Mozilla

- SpiderMonkey JavaScript Engine



MDN > Mozilla > Projects > SpiderMonkey

LANGUAGES EDIT ⚙️ ⚙️



## SpiderMonkey

### SEE ALSO

[SpiderMonkey](#)

### References:

- ▶ [JSAPI reference](#)
- ▶ [Debugger-API](#)

### Guides:

- ▶ [General](#)
- ▶ [SpiderMonkey internals](#)

### Contributing to SpiderMonkey:

- ▶ [Getting started](#)
- ▶ [Tests](#)

### Releases:

- ▶ [Release notes](#)

### Documentation:

- ▶ [Useful lists](#)

SpiderMonkey is Mozilla's [JavaScript](#) engine written in C/C++. It is used in various Mozilla products, including Firefox, and is available under the MPL2.

SpiderMonkey 38 is the most recent standalone source code release. It is largely the same engine that shipped with Firefox 38 (ESR). Full source code is available here: <https://people.mozilla.org/~sstangl/mozjs-38.2.1.rc0.tar.bz2>

The next release will be SpiderMonkey 45.

## Guides

### Building

#### [SpiderMonkey Build Documentation](#)

How to get SpiderMonkey source code, build it, and run the test suite.

### Using SpiderMonkey

## Reference

### [JS API Reference](#)

SpiderMonkey API reference.

### [JS Debugger API Reference](#)

API reference for the Debugger object introduced in SpiderMonkey 1.8.6, which corresponds to Gecko 8.0 (Firefox 8.0 / Thunderbird 8.0 / SeaMonkey 2.5).

# Apple

- Objective-C. LLVM.

The screenshot shows a web page from the Mac Developer Library. The header includes the Apple Developer logo and the title "LLVM Compiler Overview". The main content area has a dark blue header with the title "LLVM Compiler Overview". Below this, there is a large section titled "LLVM Compiler Overview" with a dark blue background. The text discusses the LLVM compiler's history, its unique approach to building compiler technologies as libraries, and its integration with Xcode and the Clang front end. It also highlights the flexibility of the library-based approach and the features available to Xcode due to the adoption of LLVM.

Mac Developer Library

Developer

LLVM Compiler Overview

## LLVM Compiler Overview

The LLVM compiler is the next-generation compiler, introduced in Xcode 3.2 for Snow Leopard, based on the open source [LLVM.org](#) project. The LLVM.org project employs a unique approach of building compiler technologies as a set of libraries. Capable of working together or independently, these libraries enable rapid innovation and the ability to attack problems never before solved by compilers. Multiple technology groups within Apple are active contributors within the LLVM.org community, and they use LLVM technology to make Apple platforms faster and more secure.

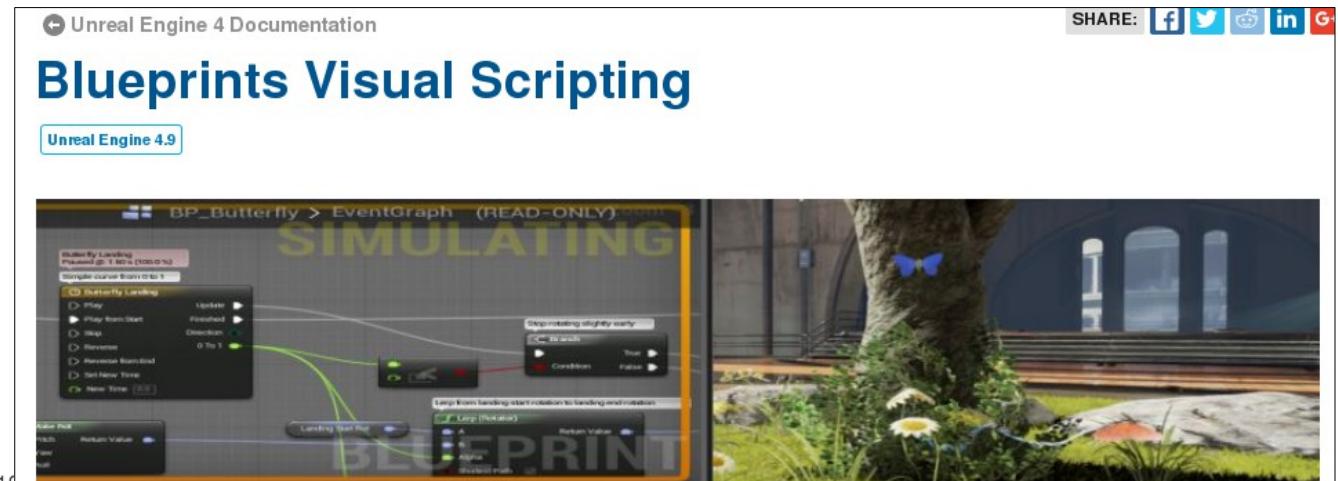
In Xcode, the LLVM compiler uses the Clang front end (a C-based languages project on LLVM.org) to parse source code and turn it into an interim format. Then the LLVM code generation layer (back end) turns that interim format into final machine code. Xcode also includes the LLVM GCC compiler, which uses the GCC compiler front end for maximum compatibility, and the LLVM back end, which takes advantage of LLVM's advanced code generator. This shows the flexibility of a library-based approach to compiler development. There are many other features, such as link-time optimization, more detailed diagnostic information, and even static analysis, that are made available to Xcode due to the adoption of LLVM.

## About Objective-C

Objective-C is the primary programming language you use when writing software for OS X and iOS. It's a superset of the C programming language and provides object-oriented capabilities and a dynamic runtime. Objective-C inherits the syntax, primitive types, and flow control statements of C and adds syntax for defining classes and methods. It also adds language-level support for object graph management and object literals while providing dynamic typing and binding, deferring many responsibilities until runtime.

# Games

- Unreal Engine: Blueprints Scripting



## QuakeC

QuakeC is an [interpreted language](#) developed in 1994 by id Software to program parts of the [video game Quake](#). A programmer is able to customize Quake to great extent by changing weapons, changing game logic and physics, and programming scenarios. It can be used to control many aspects of parts of the AI, triggers, or changes in the level. The only game engine to use QuakeC. Following engines modules for customization written in C and C++ from id Tech 4 on.

The **Blueprints Visual Scripting** system in Unreal Engine is a complete gameplay scripting system based on the concept of using a node-based interface to create gameplay elements from within Unreal Editor. This system is extremely flexible and powerful as it provides the ability for designers to use virtually the full range of concepts and tools generally only available to programmers.

Through the use of Blueprints, designers can prototype, implement, or modify virtually any gameplay element, such as:

Type static, strong  
discipline

### Major implementations

Quake C Compiler, FastQCC, QCCx, GMQCC

### Influenced by

C

### Contents

Overview

Limitations

# Games (continued)

- Osiris (Larian), Papyrus (Bethesda)

## Introduction

Osiris is a mostly declarative programming language, similar to Prolog<sup>[1]</sup>. In programming, the language itself is quite simple and only has a few constructs in progress; there are 879 Osiris APIs at the time of writing). These APIs are used to query data (queries), and to change the game state (calls).

If you are familiar with databases, you can think of an Osiris program as defining tables and rows in reaction to other database entries getting modified. Additionally, the code can change the game state. Note that the tables are actually referred to as "databases" in Osiris, so below, defining a table is generally referred to as "defining a fact", although "defining a database" or "defining a table" are also common terms.

Here is an example Osiris code fragment:

```
INIT
    DB_MyPrefix_Fruit("Apple");           // Defines the database
a fact consisting of the string "Apple" to it.
    DB_MyPrefix_Fruit("Pear");           // Adds another fact to
    DB_MyPrefix_Fruit("Banana");         // And one more.
```

## Papyrus

Papyrus is an object-oriented scripting language developed by Bethesda Softworks for use in the Elder Scrolls V: Skyrim game engine. Papyrus first appeared in Elder Scrolls V: Skyrim and has since been utilized in Fallout 4 and The Elder Scrolls Online. Papyrus is an evolution of ObScript,<sup>[2]</sup> the previous scripting language used in Fallout 3 and Fallout: New Vegas, with additional functionality and flexibility along with completely new syntax and workflow.

Papyrus works by receiving in-game Events and sending Function Calls. Papyrus is the driver for all game events and allows tracking of variables based on player interactions. Papyrus also handles the use of triggers, player terminals and book interactions, as well as special interactions through dialogue with characters.

Unlike the previous language, Papyrus scripts must be written in a separate text editor and compiled using the Creation Kit to be used in-game. Plugins for Notepad++ and Atom allow for use of Papyrus syntax.

Is the scripting language for the hot new game really “similar to Prolog”? What does that mean?

# Compilers and Interpreters

- Back End Optimization, Chips, etc.
  - Intel, AMD, nVidia, Green Hills, etc.
- Platform Vendors
  - Apple, Oracle, etc.
- Tooling, IDEs
  - Microsoft, Google, etc.
- Domain-Specific Languages
  - Photoshop, Game Studies, MATLAB, SQL, Wolfram Alpha, etc.

# Surprise: Postscript /PDF

## ^ The language



PostScript is a Turing-complete programming language, belonging to the [concatenative](#) group. Typically, PostScript programs are not produced by humans, but by other programs. However, it is possible to write computer programs in PostScript just like any other programming language.<sup>[5]</sup>

PostScript is an [interpreted, stack-based](#) language similar to [Forth](#) but with strong dynamic [typing](#), data structures inspired by those found in [Lisp](#), [scoped memory](#) and, since language level 2, [garbage collection](#). The language syntax uses [reverse Polish notation](#), which makes the order of operations unambiguous, but reading a program requires some practice, because one has to keep the layout of the [stack](#) in mind. Most *operators* (what other languages term *functions*) take their arguments from the stack, and place their results onto the stack. [Literals](#) (for example, numbers) have the effect of placing a copy of themselves on the stack. Sophisticated data structures can be built on the [array](#) and [dictionary](#) types, but cannot be declared to the type system, which sees them all only as arrays and dictionaries, so any further typing discipline to be applied to such user-defined "types" is left to the code that implements them.

**Portable Document Format (PDF)**, standardized as [ISO 32000](#), is a [file format](#) developed by [Adobe](#) in 1992 to present [documents](#), including text formatting and images, in a manner independent of [application software](#), [hardware](#), and [operating systems](#).<sup>[2][3]</sup> Based on the [PostScript](#) language, each PDF file encapsulates a complete description of a fixed-layout flat document, including the text, [fonts](#),

# Surprise: Postscript /PDF

## ^ The language



PostScript is a Turing-complete programming language, belonging to the [concatenative](#) group. Typically, PostScript programs are not produced by humans, but by other programs. However, it is possible to write computer programs in PostScript just like any other programming language.<sup>[5]</sup>

PostScript is an [interpreted](#), [stack-based](#) language similar to [Forth](#) but with strong dynamic [typing](#), data structures inspired by those found in [Lisp](#), [scoped memory](#) and, since language level 2, [garbage collection](#). The language syntax uses [reverse Polish notation](#), where one has to keep the operators on the stack, and place the values to be operated on the stack. Sophisticated systems, which see what "types" is left to the interpreter, can even do type inference.

Any PDF reader, a cell phone, a game console, a car or bus: they all contain *Interpreters*.

### Portable Document Format

developed by [Adobe](#) in 1992 to present [documents](#), including [text](#) [formatting](#) and [images](#), in a manner independent of [application software](#), [hardware](#), and [operating systems](#).<sup>[2][3]</sup> Based on the [PostScript](#) language, each PDF file encapsulates a complete description of a fixed-layout flat document, including the text, [fonts](#),

# Who Cares?

- The *computer* is unique among “machines” (e.g., lever, pulley, etc.) in that it magnifies our mental force rather than our physical force.
  - Computers can assist with decision making, model and predict outcomes, etc.
- **Programming Languages** are the mechanism for communicating with and commanding the only tool available that magnifies your mind.

# Plus Work, Double-Plus Easy

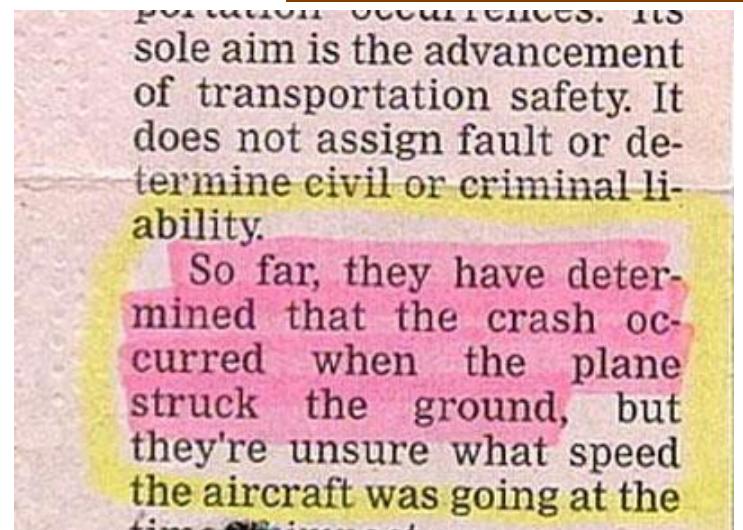
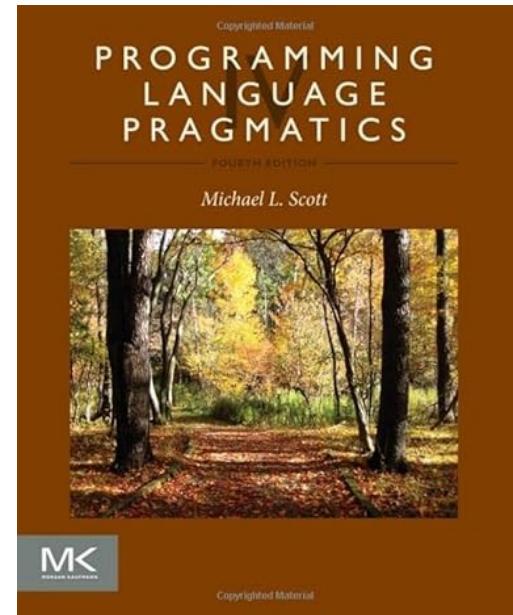
- Unhappiness is related to unrealized desires or unmet expectations
- At some schools, PL is arguably the most difficult CS course
- Principles of Programming Languages (CSCI 2320) is approachable
  - Significant theoretical component
  - Significant programming component
  - Generous curve and opportunities for success

# Course Structure

- Course has **theoretical** and **practical** aspects
  - Best of both worlds!
- Need both in programming languages!
- **Reading = both**
  - Many external and optional readings
- **Review Sets = theory**
  - Not graded, practice problems for exams
- **Programming Assignments = practice**
  - Electronic hand-in
- Strict deadlines but 3x “late days”
  - Ask me why ...

# Resources

- Textbook
  - Programming Language Pragmatics
  - Michael L. Scott
- Video Guides
- Free Online Materials
  - Udacity CS 262
- Optional Readings



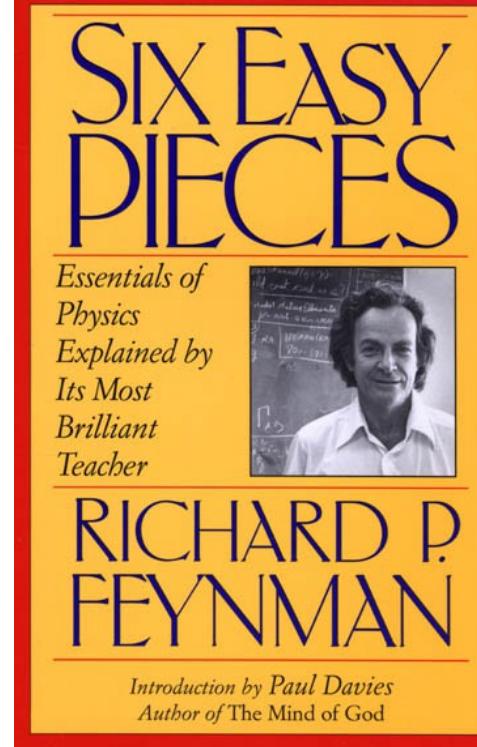
# Academic Honesty

- Don't use work from uncited sources
  - Including old code
- We often use plagiarism detection software
- Class discussion later
  - ChatGPT: allowed? No?

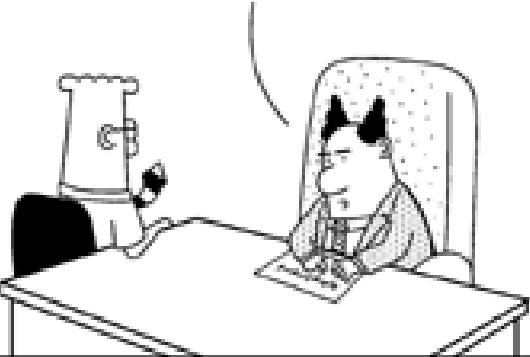


# LDI Course Project

- A big project: an Interpreter!
- ... in four easy parts
- You may optionally work in pairs.

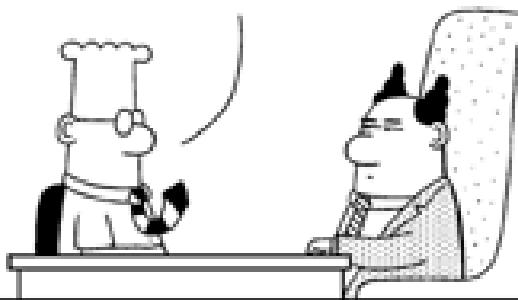


WHAT DOES MFU2  
MEAN ON YOUR  
TIMELINE?



[www.dilbert.com](http://www.dilbert.com)

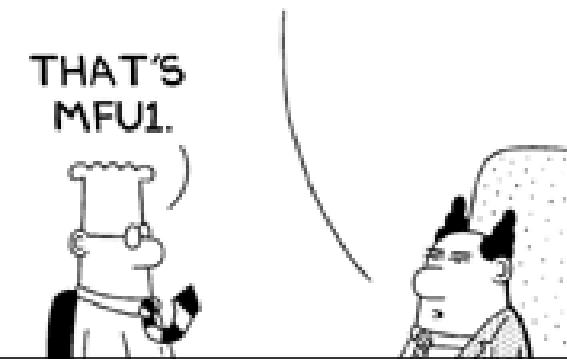
THAT'S MANAGEMENT  
FOUL-UP NUMBER TWO.  
IT USUALLY HAPPENS  
AROUND THE THIRD  
WEEK.



1-30-06 © 2006 Scott Adams, Inc./Dist. by UFS, Inc.

WE DON'T ANTICIPATE  
ANY MANAGEMENT  
MISTAKES.

THAT'S  
MFU1.



# “Explaining Unicorns & Dragons”

- Visual Studio, JVM, Exceptions, Memory, Debugging, Linking, Shared Libraries, ...

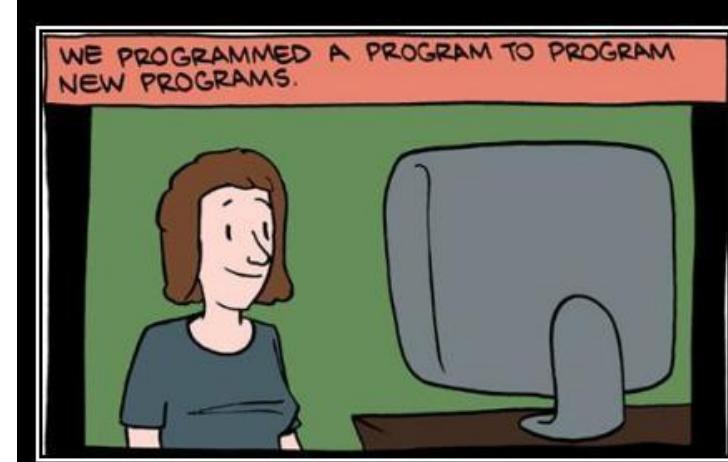


# How are Languages Implemented?

- Two major strategies:
  - Interpreters (take source code and run it)
  - Compilers (translate source code, run result)
  - Distinctions blurring (e.g., just-in-time compiler)
- Interpreters run programs “as is”
  - Little or no preprocessing
- Compilers do extensive preprocessing
  - Most implementations use compilers

# (Short) History of High-Level Languages

- 1953 IBM develops the 701 “Defense Calculator”
  - 1952, US formally ends occupation of Japan
  - 1954, Brown v. Board of Education of Topeka, Kansas
- All programming done in assembly
- Problem: Software costs exceeded hardware costs!
- John Backus: “Speedcoding”
  - An interpreter
  - Ran 10-20 times slower than hand-written assembly





# FORTRAN I

- 1954 IBM develops the 704
- John Backus
  - Idea: translate high-level code to assembly
  - Many thought this impossible
- Set in ~1961
- 1954-7 FORTRAN I project
- By 1958, >50% of all software is in FORTRAN
- Cut development time dramatically
  - (2 weeks → 2 hours)

# FORTRAN I

- The first **compiler**
  - Produced code almost as good as hand-written
  - Huge impact on computer science
- Led to an enormous body of theoretical work
- Modern compilers keep the outlines of FORTRAN I



Grace Hopper



# Changeups and Trivia

- “[Professors who] deliberately and consistently interspersed their lectures with ... some other form of deliberate break ... usually commanded a better attention span from the class, and these deliberate variations had the effect of postponing or even eliminating the occurrence of an attention break”

[ Johnstone and Percival. *Attention breaks in lectures*. Education in Chemistry, 13. 49-50, 1976. ]

[ Middendorf and Kalish. *The “Change-up” in Lectures*. TRC Newsletter, 8:1 (Fall 1996). ]

# Real-World Languages

- This Indo-European language is associated with South Asian Muslims and is the lingua franca of Pakistan. It developed from Persian, Arabic and Turkic influences over about 900 years. Poetry in this language is particularly famed, and is a reported favorite of former US President Barack Obama.
- Example: **السلام عليكم**

# Interpreters

Lexical Analysis

Parsing

Semantic Analysis

Optimization (optional)

Interpret The Program

# Compilers

Lexical Analysis

Parsing

Semantic Analysis

Optimization (optional)

Generate Machine Code

Run that Machine Code

The first three may benefit from an analogy to how humans comprehend English.

# Lexical Analysis

- First step: recognize words.
  - Smallest unit above letters

This is a sentence.

- Note the
  - Capital “T” (start of sentence symbol)
  - Blank “ ” (word separator)
  - Period “.” (end of sentence symbol)



# More Lexical Analysis

- Lexical analysis is not trivial. Consider:  
**How d'you break “this” up?**
- Plus, programming languages are typically more cryptic than English:  
**\*p->f += -.12345e-6**



# And More Lexical Analysis

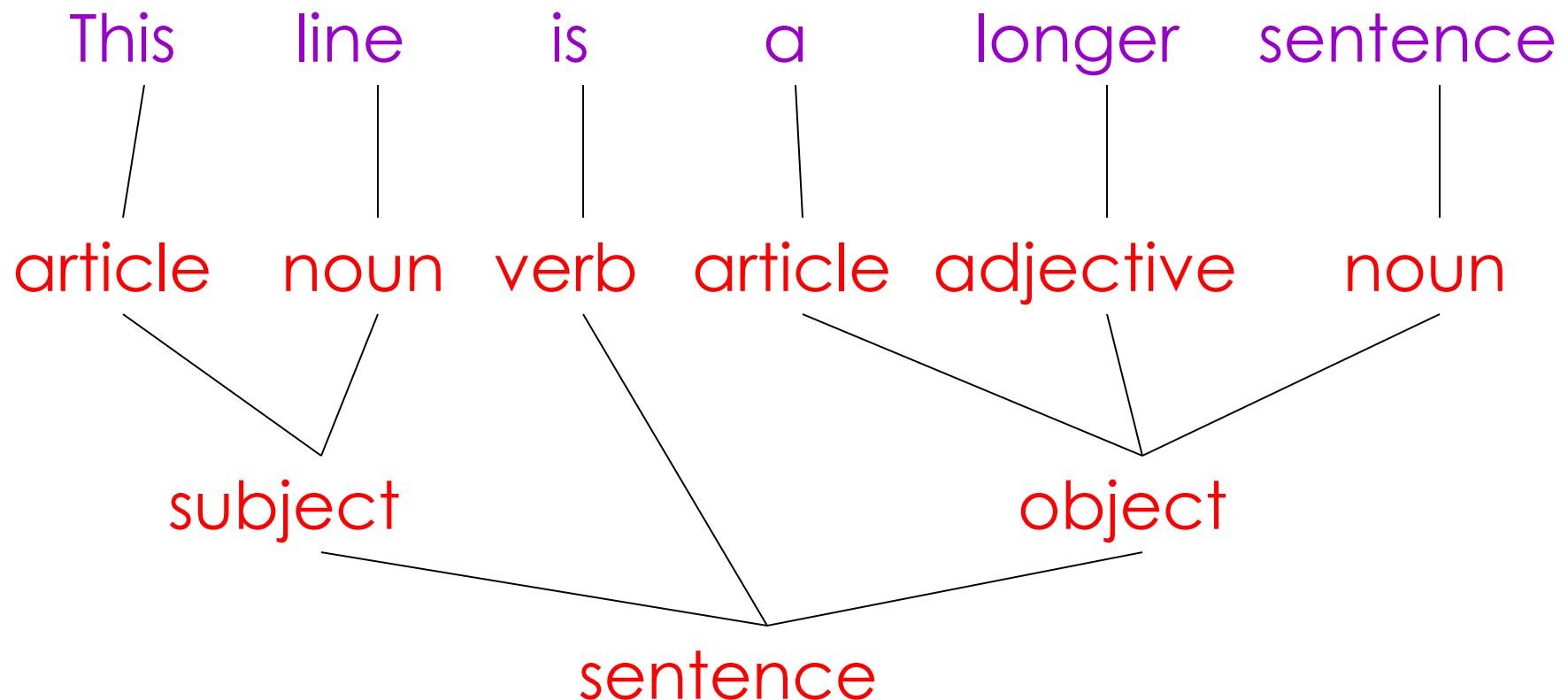
- Lexical analyzer divides program text into “words” or tokens  
**if x == y then z = 1; else z = 2;**
- Broken up:  
**if, x, ==, y, then, z, =, 1, ;, else, z, =, 2, ;**

# Parsing

- Once words are understood, the next step is to understand sentence structure
- Parsing = Diagramming Sentences
  - The diagram is a tree
  - Often annotated with additional information



# Diagramming a Sentence

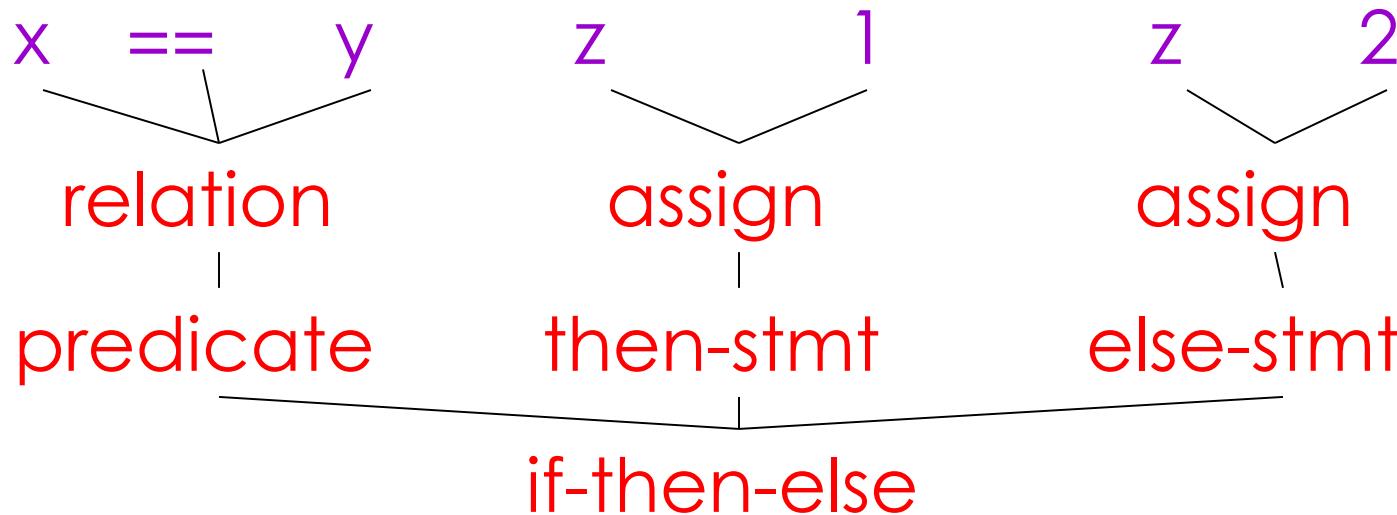


# Parsing Programs

- Parsing program expressions is the same
- Consider:

if x == y then z = 1; else z = 2;

- Diagrammed:



# Semantic Analysis

- Once sentence structure is understood, we can try to understand “meaning”
  - But meaning is **too hard for compilers**
- Compilers perform limited analysis to catch inconsistencies: **reject bad programs early!**
- Some do more analysis to improve the performance of the program

# Semantic Analysis in English

- Example:

Arya said Sansa left her direwolf at home.

What does “her” refer to? Arya or Sansa?



- Even worse:

No One said No One left her mask at home.

How many “No One”s are there?

Which one left the mask?

It's  
context-  
sensitive!

# Semantic Analysis in Programming

- Programming languages define strict rules to avoid such ambiguities
- This C++ code prints “4”; the inner definition is used

```
{  
    int Sydney = 3;  
    {  
        int Sydney = 4;  
        cout << Sydney;  
    }  
}
```





# Finding Disagreements

- Compilers perform many semantic checks besides variable bindings
- *Potential* example:  
**Gregory House left her cane at home.**
- This may be a “type mismatch” between **her** and **Gregory House** (depending on gender)
  - If “him” should be used for this Gregory House
  - Just as in real life, this is context-sensitive!
    - “Her” can be correct for some Gregory Houses.

# Optimization

- No strong counterpart in English, but akin to editing (cf. poems, short stories)
- Automatically modify programs so that they
  - Run faster
  - Use less memory
  - Use less energy (e.g., on your phone)
  - In general, conserve some resource

# Code Generation

- Produces assembly code (usually)
  - which is then assembled into an executable by an assembler
- A translation into another language
  - Analogous to human translation
- (Code generation in PL is not really analogous to ChatGPT. We'll cover it later!)

# Issues

- Compiling and interpreting are almost this simple, but there are **many pitfalls**.
- Example: How are bad programs handled?
- Language design has big impact on compiler
  - Determines what is easy and hard to compile
  - Course theme: **trade-offs in language design**



# Languages Today

- The overall structure of almost every compiler & interpreter follows our outline
- The proportions have changed since FORTRAN
  - Early: lexing, parsing most complex, expensive
  - Today: optimization dominates all other phases, lexing and parsing are cheap
  - Thus: this course puts no emphasis on ancient parsing optimizations (e.g., LL, LALR)

# Trends in Languages

- Optimization for speed is less interesting. But:
  - scientific programs
  - advanced processors (Digital Signal Processors, advanced speculative architectures)
  - small devices where speed = longer battery life
- Ideas we'll discuss are used for **improving code reliability**:
  - memory safety
  - detecting concurrency errors (data races)
  - type safety
  - automatic memory management
  - ...

# Why Study Prog. Languages?

- Prepare for many good jobs
- Increase capacity of expression
- Improve understanding of program behavior
  - Know how things work “under the hood”
- Increase ability to learn new languages
- Learn to build a large and reliable system
- See many basic CS concepts at work
- Computers are the only tools that increase cognitive power, so learn to control them

# What Will You Do In This Class?

- Reading (textbook, videos, outside sources)
- Learn about different kinds of languages
  - Imperative vs. Functional vs. Object-Oriented
  - Static typing vs. Dynamic typing
  - etc.
- Gain exposure to new languages (ML, Cool)
- Write an interpreter!

# What Is This?

A lungo il mio cuore di tali ricordi ha voluto colmarsi!

Come un vaso in cui le rose sono state dissetate:

Puoi romperlo, puoi distruggere il vaso se lo vuoi,

Ma il profumo delle rose sarà sempre tutt'intorno.

Mon coeur est brûlant rempli de tels souvenirs

Comme un vase dans lequel des roses ont été distillées:

Tu peux le briser, tu peux détruire le vase si tu le désires,

Mais la senteur des roses sera toujours là.

Długo, długo moje serce przepelcone było takimi wspomnieniami!

Były jak waza, w której kiedyś róże destylowały:

Możesz sprawić by pekła, możesz gruchotać waze jeśli chcesz,

Ale zapach róż bedzie wciąż czuć dookoła.

Lang, lang soll die Erinnerung in meinem Herzen klingen!

Gleich einer Vase, drin Rosen sich einst tränkten:

Lass sie zerbrechen, lass sie zerspringen,

Der Duft der Rose bleibt immer hängen.

Muito, muito tempo seja meu coração preenchido com tais lembranças!

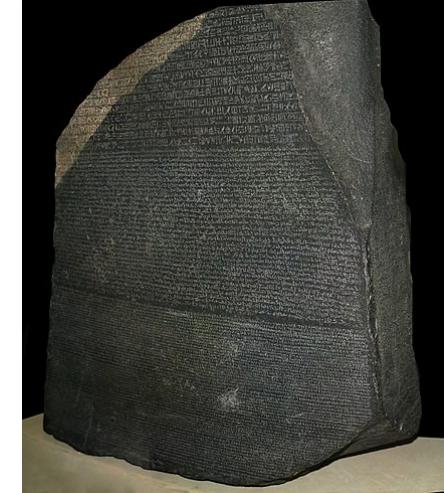
Tal qual o vaso onde rosas foram uma vez destiladas:

Pode quebrar, pode estilhaçar o vaso se o desejas,

Mas perdurará para sempre o aroma das rosas perfumadas.

# The Rosetta Stone

- The first programming assignment involves writing the same simple (50-75 line) program in two languages:
  - Ocaml and Cool (with Ruby, Python, JavaScript, Haskell and C as Extra Credit)
- PA1c, due Tue Jan 30, requires you to write the program in one language
- PA1, due subsequent Thursday, requires both



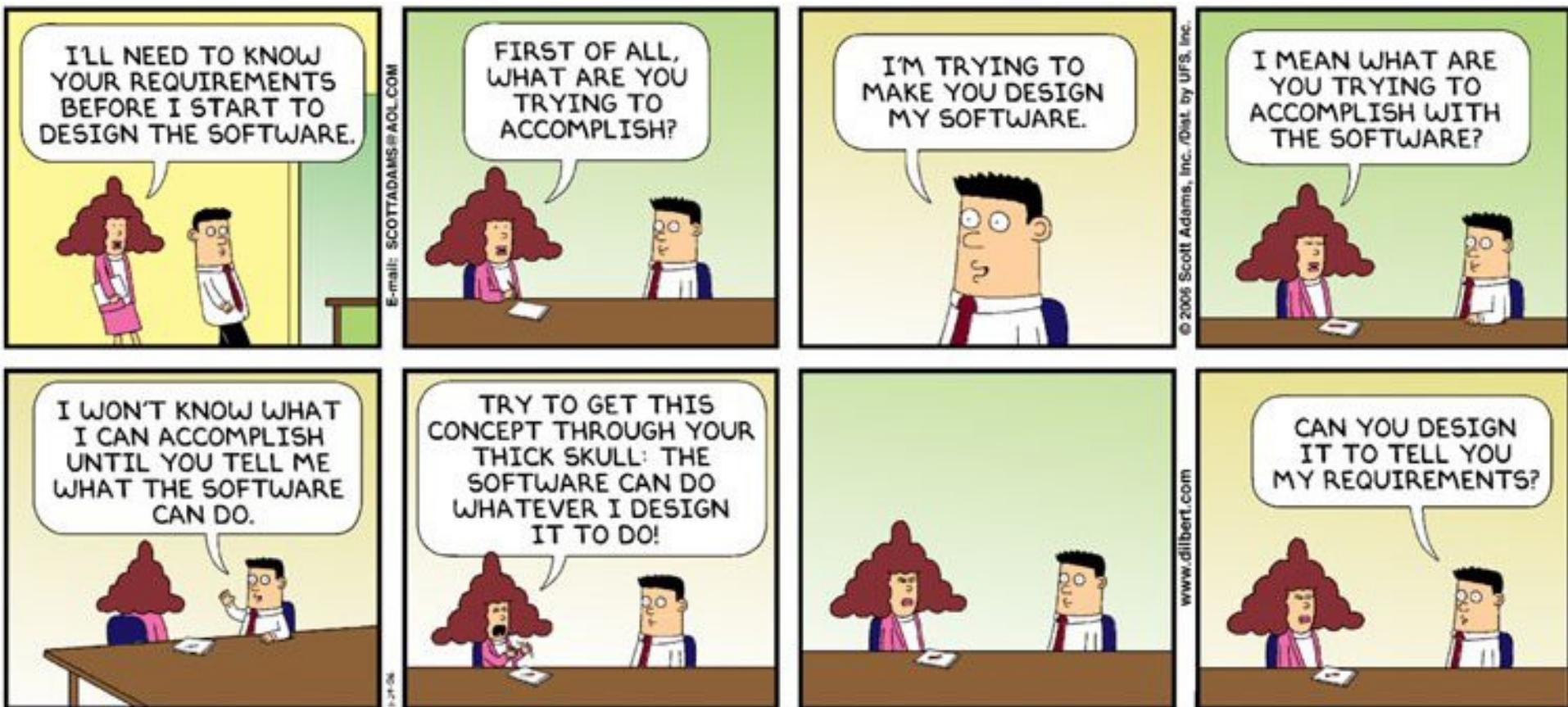
Long, long be my heart with such memories fill'd!  
Like the vase in which roses have once been distill'd:  
You may break, you may shatter the vase if you will,  
But the scent of the roses will hang round it still.  
- Thomas Moore (Irish poet, 1779-1852)

# Partial Automated Grading

- You should think about the project on your own, write your own (local) test cases, and then submit to the grading server
  - Limited submissions per day
  - Discourages “guess and check”
- Continuous Integration Testing, Alpha & Beta Testing
- Ecological Validity
- Job Interviews (LeetCode, HackerRank, etc.)
- Grade and Time Control

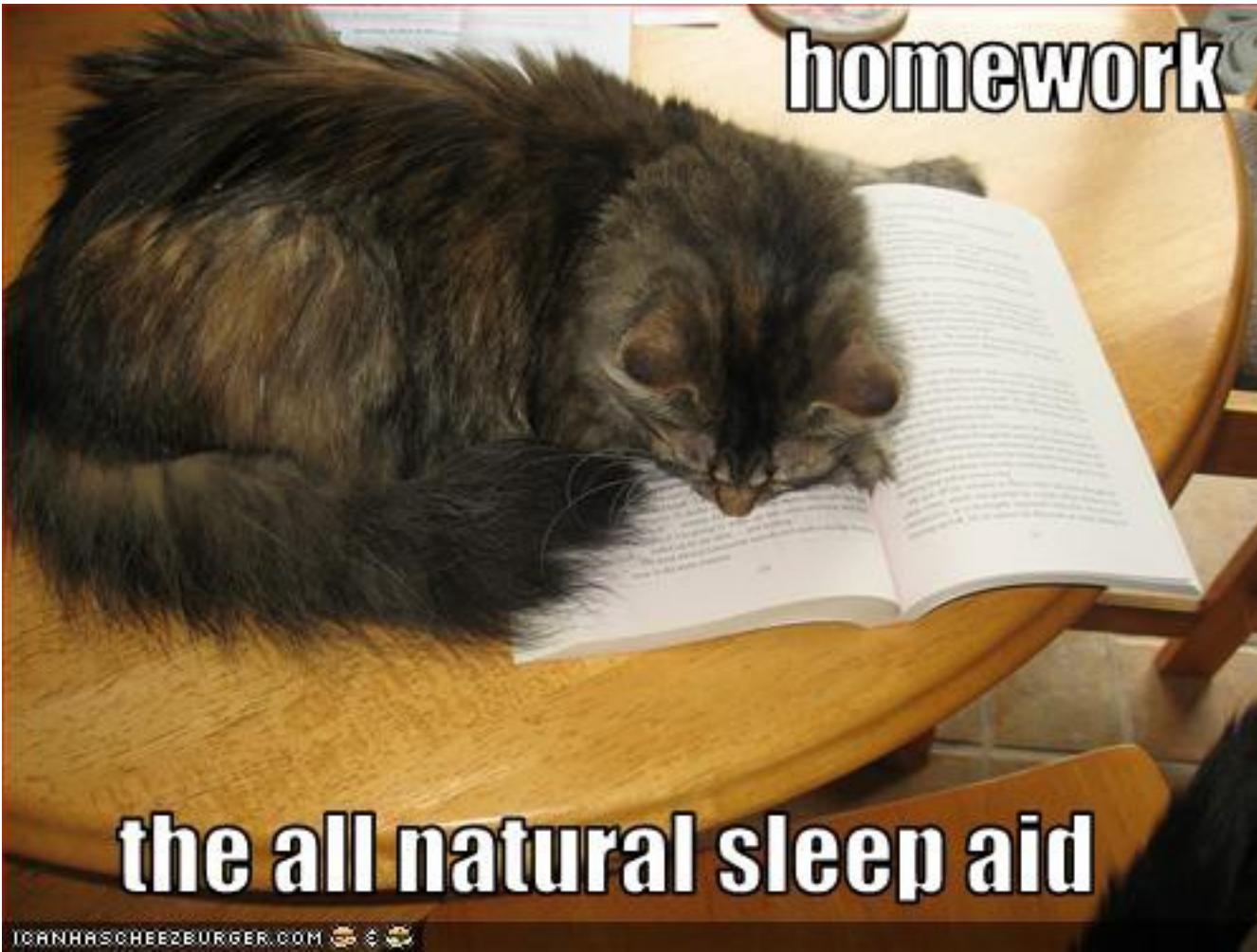
# Live Submission Demo

- Let's visit the automated submission website



# Start The Homework Now

- We can help you! (Also: video guides ...)



# Homework

- Scott Book reading (for Tuesday)
- Get started on PA1c (due in 7 days)

## Questions?

