Programming Languages

CS242
Lecture 1
Course Staff

Alex Aiken

Timothy Gu

Rohan Yadav

Jerry Chen

Xiaoyuan Ni
Important Info

See cs242.stanford.edu for other important course information
CS242: Programming Languages

• All things programming languages!

• Different families of programming languages
  • Including some you may have not seen before
  • In theory and practice

• Core ideas
  • Semantics
  • Type systems
  • Program analysis
  • Formal verification

• Core features
  • Continuations
  • Monads
  • Concurrency
  • Parallelism
Course Prerequisites

• Theory: CS 103
  • First-order logic, induction, discrete math

• Systems: CS 107 + 110

• Comfortable with programming
  • And learning new programming systems
Course Structure

• Homeworks
  • 8 programming assignments
  • Roughly weekly, starting second week (56%)
  • Expect ~10 hours/week

• Readings

• Two exams
  • In-class midterm (19%)
  • Cumulative final (25%)
Takeaways

• Appreciation for programming languages as a technical field
  • The problems, the values, the “culture”

• Becoming a better programmer
  • Learn to think systematically about programming tools

• Understand the future
  • What we can do, what we can’t do, what we will likely be able to do

• Basics of active research areas
  • Preparation for research

• Thanks to Will Crichton for many of the slides in this lecture!
The Big Picture: Language Goals

Productivity

Python

ML, Haskell

Java, C++

Safety

Coq

Rust

Performance

Matlab

C

Alex Aiken      CS 242     Lecture 1
Language Goals

• Every programming language has as goals
  • Performance
  • Productivity
  • Safety

• But there are tradeoffs

• And different designs make different choices
  • One of the reasons we have so many programming languages
Programming Languages are Complex Tools

Product details
Publisher: Addison-Wesley Professional; 4th edition (May)
Language: English
Paperback: 1376 pages
Item Weight: 2.73 pounds

Product details
Publisher: Wiley; 1st edition (July 8, 2014)
Language: English
Paperback: 1152 pages
ISBN-10: 1118907442
Item Weight: 6.04 pounds

Product details
Publisher: McGraw-Hill Education;
Language: English
Paperback: 1248 pages
ISBN-10: 1260440230
Item Weight: 3.86 pounds

Product details
ASIN: B07CQPHC1N
Publication date: April 27, 2018
Language: English
File size: 70555 KB
Text-to-Speech: Enabled
Enhanced typesetting: Enabled
X-Ray: Not Enabled
Word Wise: Not Enabled
Print length: 377 pages
Developers Are Reading, Not Writing


<table>
<thead>
<tr>
<th></th>
<th>Comprehension</th>
<th>Navigation</th>
<th>Editing</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>57.62%</td>
<td>23.96%</td>
<td>5.02%</td>
<td>13.40%</td>
</tr>
</tbody>
</table>

Complexity
move_if_noexcept obtains an rvalue reference to its argument if its move constructor does not throw exceptions or if there is no copy constructor (move-only type), otherwise obtains an lvalue reference to its argument. It is typically used to combine move semantics with strong exception guarantee.

**Parameters**

x - the object to be moved or copied

**Return value**

std::move(x) or x, depending on exception guarantees.

**Notes**

This is used, for example, by std::vector::resize, which may have to allocate new storage and then move or copy elements from old storage to new storage. If an exception occurs during this operation, std::vector::resize undoes everything it did to this point, which is only possible if std::move_if_noexcept was used to decide whether to use move construction or copy construction. (unless copy constructor is not available, in which case move constructor is used either way and the strong exception guarantee may be waived)

**Example**

```cpp
#include <iostream>
#include <utility>

template< class T >
typedef std::conditional<
    !std::is_nothrow_move_constructible<T>::value && std::is_copy_constructible<T>::value,
    const T&,
    T&
> ::type move_if_noexcept(T& x) noexcept;

template< class T >
constexpr typename std::conditional<
    !std::is_nothrow_move_constructible<T>::value && std::is_copy_constructible<T>::value,
    const T&,
    T&
> ::type move_if_noexcept(T& x) noexcept;
```
6.7.3.1 Formal definition of restrict

1. Let $D$ be a declaration of an ordinary identifier that provides a means of designating an object $P$ as a restrict-qualified pointer to type $T$.

2. If $D$ appears inside a block and does not have storage class `extern`, let $B$ denote the block. If $D$ appears in the list of parameter declarations of a function definition, let $B$ denote the associated block. Otherwise, let $B$ denote the block of `main` (or the block of whatever function is called at program startup in a freestanding environment).

3. In what follows, a pointer expression $E$ is said to be based on object $P$ if (at some sequence point in the execution of $B$ prior to the evaluation of $E$) modifying $P$ to point to a copy of the array object into which it formerly pointed would change the value of $E$.\(^{119}\) Note that "based" is defined only for expressions with pointer types.

4. During each execution of $B$, let $L$ be any lvalue that has $&L$ based on $P$. If $L$ is used to access the value of the object $X$ that it designates, and $X$ is also modified (by any means), then the following requirements apply: $T$ shall not be const-qualified. Every other lvalue used to access the value of $X$ shall also have its address based on $P$. Every access that modifies $X$ shall be considered also to modify $P$, for the purposes of this subclause. If $P$ is assigned the value of a pointer expression $E$ that is based on another restricted pointer object $P_2$, associated with block $B_2$, then either the execution of $B_2$ shall begin before the execution of $B$, or the execution of $B_2$ shall end prior to the assignment. If these requirements are not met, then the behavior is undefined.

5. Here an execution of $B$ means that portion of the execution of the program that would correspond to the lifetime of an object with scalar type and automatic storage duration
Intel CPU Trends
(sources: Intel, Wikipedia, K. Olukotun)

Processor clock rate stops increasing
No further benefit from ILP

Image credit: "The free Lunch is Over" by Herb Sutter, Dr. Dobbs 2005  
CMU 15-418/618, Spring 2017
Memory usage continues to scale

<table>
<thead>
<tr>
<th>Process Name</th>
<th>Memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slack Helper</td>
<td>365.8 MB</td>
</tr>
<tr>
<td>Slack Helper</td>
<td>234.9 MB</td>
</tr>
<tr>
<td>Slack</td>
<td>91.8 MB</td>
</tr>
<tr>
<td>Slack Helper</td>
<td>89.8 MB</td>
</tr>
</tbody>
</table>

700 MB!

3 GB!!!
Huge Performance Gaps

<table>
<thead>
<tr>
<th></th>
<th>Immutable</th>
<th>Mutable</th>
<th>Double Only</th>
<th>No Objects</th>
<th>In C</th>
<th>Transposed</th>
<th>Tiled</th>
<th>Vectorized</th>
<th>BLAS MxM</th>
<th>BLAS Parallel</th>
</tr>
</thead>
<tbody>
<tr>
<td>ms</td>
<td>17,094,152</td>
<td>77,826</td>
<td>32,800</td>
<td>15,306</td>
<td>7,530</td>
<td>2,275</td>
<td>1,388</td>
<td>511</td>
<td>196</td>
<td>58</td>
</tr>
</tbody>
</table>

- 219.7x
- 2.4x
- 2.2x
- 3.4x
- 2.8x
- 3.5x
- 219.7x
- 522x
- 1117x
- 2271x
- 7514x
- 12316x
- 33453x
- 87042x
- 296260x

Cycles/OP | 8,358 | 38 | 16 | 7 | 4 | 1 | 1/2 | 1/5 | 1/11 | 1/36 |

Saman Amarasinghe, MIT 6.172 “Performance Engineering”. 2009
And Persistent Correctness Gaps

“The majority of vulnerabilities fixed and with a CVE assigned are caused by developers inadvertently inserting memory corruption bugs into their C and C++ code.”

Microsoft Security Response Center. “A proactive approach to more secure code.” 2019
Recap

1. Developers mostly comprehend/debug code
2. Software ecosystems are complex at all scales
3. The hardware is changing and software often doesn’t exploit it well
4. Bugs aren’t getting any better
The Long View
Alan Turing
“On Computable Numbers” 1937

Alonzo Church
“A set of postulates for the foundation of logic”, 1932
Turing machines

Lambda calculus

\[
\begin{align*}
[x \rightarrow y] x &= y \\
[x \rightarrow y] z &= z \\
[x \rightarrow y] \lambda z \cdot x &= \lambda z \cdot y \\
[x \rightarrow y] \lambda y \cdot x y &= \lambda y' \cdot y y' \\
[x \rightarrow y] x (\lambda x \cdot x) &= y (\lambda x \cdot x)
\end{align*}
\]
On compilers:
At that time, the Establishment told us that a computer could not write a program; it was totally impossible; that all that computers could do was arithmetic; that it had none of the imagination and dexterity of a human being. I kept trying to explain that we were wrapping up the human being’s dexterity in the program that he wrote.

On assembly:
I think I spent 20 years fighting the “Establishment.” In the early years of programming languages, the most frequent phrase we heard was that the only way to program a computer was in octal. Of course, a few years later a few people admitted that maybe you could use assembly language.
Prior to FORTRAN, most source language operations were not machine operations. Large inefficiencies in looping and computing addresses were masked by time spent in floating point subroutines.

The advent of the 704 with built-in floating point and indexing radically altered the situation. ...It increased the problem of generating efficient programs by an order of magnitude by speeding up floating point operations by a factor of ten and thereby leaving inefficiencies nowhere to hide. This caused us to regard the design of the translator as the real challenge, not the simple task of designing the language.
Algol introduced into programming languages such terms as type, declaration, identifier, for statement, while, if then else, switch, the begin end delimiters, block, call by value and call by name, typed procedures, declaration scope, dynamic arrays, side effects, global and local variables.

Algol was strongly derived from FORTRAN and its contemporaries. The logic, arithmetic and data organizations were close to those then being designed into real computers. Certain simple generalizations of computer instructions such as switch, for statement, and if statements were included because their semantics and computer processing were straight-forward consequences of single statement processing.
C fits firmly in the traditional procedural family typified by Fortran and Algol 60. It is ‘close to the machine’ in that the abstractions it introduces are readily grounded in the concrete data types and operations supplied by conventional computers.

The most important [historical accident] has been the tolerance of C compilers to errors in type. C evolved from typeless languages. It did not suddenly appear to its earliest users as an entirely new language with its own rules; instead we continually had to adapt existing programs as the language developed, and make allowance for an existing body of code.

Ken Thompson,
Dennis Ritchie
“The Development of the C Language”
1993
Our first main task was to carry out resonance absorption calculations related to the construction of Norway's first nuclear reactor. Monte Carlo simulation methods were successfully introduced instead in 1949-1950. The necessity of using simulation and the need of a language for system description was the direct stimulus for SIMULA.

When writing simulation programs we had observed that processes often shared a number of common properties, both in data attributes and actions, but were structurally different in other respects so that they had to be described by separate declarations. Such partial similarity fairly often applied to processes in different simulation models, indicating that programming effort could be saved by somehow preprogramming the common properties.
Simula’s class concept allowed me to map my application concepts into the language constructs in a direct way. The way Simula classes can act as coroutines made the inherent concurrency of my application easy to express.

The implementation of Simula, however, did not scale in the same way and as a result the whole project came close to disaster. The cost arose from several language features and their interactions: run-time type checking, guaranteed initialization of variables, concurrency support, and garbage collection of both user-allocated objects and procedure activation records.
The team originally considered using C++, but rejected it for several reasons. They decided that C++’s complexity led to developer errors. The language's lack of garbage collection meant that programmers had to manually manage system memory, a challenging and error-prone task. The team also worried about the C++ language's lack of portable facilities for security, distributed programming, and threading. Finally, they wanted a platform that would port easily to all types of devices.

— Wikipedia
My initial goal for Python was to serve as a second language for people who were C or C++ programmers, but who had work where writing a C program was just not effective.

Maybe it was something you'd do only once. It was the sort of thing you'd prefer to write a shell script for, but when you got into the writing details, you found that the shell was not the ideal language—you needed more data structures, more namespaces, or maybe more performance. The first sound bite I had for Python was, "Bridge the gap between the shell and C."
Turing languages

1957 — FORTRAN
1959 — ALGOL
1962 — SIMULA
1972 — C
1979 — C++
1991 — Python
1995 — Java
My own research in artificial intelligence [in 1958]... involved representing information about the world by sentences in a suitable formal language and a reasoning program that would decide what to do by making logical inferences. Representing sentences by list structure seemed appropriate and a list-processing language also seemed appropriate for programming the operations involved in deduction.

...One needs a notation for functions, and it seemed natural to use the λ-notation of Church (1941). I didn't understand the rest of his book, so I wasn't tempted to try to implement his more general mechanism for defining functions.
The languages people use to communicate with computers differ in their intended aptitudes, towards either a particular application area, or a particular phase of computer use (high level programming, program assembly, job scheduling, etc). The question arises, do the idiosyncrasies reflect basic logical properties of the situations that are being catered for? Or are they accidents of history and personal background that may be obscuring fruitful developments?

ISWIM is an attempt at a general purpose system for describing things in terms of other things, that can be problem-oriented by appropriate choice of "primitives." A possible first step in the research program is 1700 doctoral theses called "A Correspondence between $\mathbf{x}$ and Church's $\lambda$-notation."

Peter Landin, “The Next 700 Programming Languages” 1966
The principal aims in designing ML were to make it impossible to prove non-theorems yet easy to program strategies for performing proofs.

A strategy—or recipe for proof—could be something like “induction on $f$ and $g$, followed by assuming antecedents and doing case analysis, all interleaved with simplification”. This is imprecise—analysis of what cases? what kind of induction, etc, etc.—but these in turn may well be given by further recipes, still in the same style.

Robin Milner, “A Metalanguage for Interactive Proof in LCF” 1978
The simplicity and elegance of functional programming captivated the present authors. Lazy evaluation— with its direct connection to the pure, call-by-name lambda calculus, the remarkable possibility of representing and manipulating infinite data structures, and addictively simple and beautiful implementation techniques—was like a drug.

Hudak et al.
“A History of Haskell: Being Lazy with Class” 2007
Church Languages Are Unfamiliar

Couldn't match type `k0' with `b'
because type variable `b' would escape its scope
This (rigid, skolem) type variable is bound by
the type signature for

``
groupBy :: Ord b => (a -> b) -> Set a -> Set (b, [a])
``
The following variables have types that mention k0
PL Theory is Unfamiliar (and Dense)

\[ \Sigma_0; \Gamma_1 \vdash A = D \bar{\nu} : \text{Set}_n \quad \Gamma = \Gamma_1(x : A)\Gamma_2 \]

data D \Delta : \text{Set}_n where \( c_i \Delta_i [\bar{j}_i | b_i] \in \Sigma_0 \quad \exists k. [\bar{j}_k | b_k] \neq [] \)

\[
\begin{aligned}
\Delta'_i &= \Delta_i(\bar{j}_i : \mathbb{I})[\bar{\nu} / \Delta] & \bar{q}_i &= \Delta'_i[\bar{j}_i | b_i][\bar{\nu} / \Delta] \\
\rho_i &= 1_{\Gamma_1} \uplus [c_i \bar{q}_i / x] & \rho'_i &= \rho_i \uplus 1_{\Gamma_2} \\
\Theta_i &= \text{BOUNDARY}(\bar{j}_i); \Theta \\
\Sigma_i; \Gamma_1 \Delta'_i(\Gamma_2 \rho_i) \vdash f \bar{q} \rho'_i := Q_i : C \rho'_i | \Theta_i \leadsto \Sigma_i \\
\end{aligned}
\]

\[ \Delta_{hc} = (r : \mathbb{I})(u : \mathbb{I} \rightarrow \text{Partial} r (D \bar{\nu}))(u_0 : D \bar{\nu} [r \leftrightarrow u \text{i0}]) \]

\[ \rho_{hc} = 1_{\Gamma_1} \uplus [\text{hcomp} r u u_0 / x] \]

\[ \Sigma_n; \Gamma_1 \Delta_{hc}(\Gamma_2 \rho_{hc}) \vdash f \bar{q} \rho'_{hc} := Q_{hc} : C \rho'_{hc} | (r = 1); \Theta \leadsto \Sigma_{n+1} \]

\[ \Sigma_0; \Gamma \vdash f \bar{q} := \text{case}_{\mathbb{X}} \left\{ \begin{array}{ll}
c_1 \bar{q}_1 \mapsto Q_1; & \ldots; \\
h_{\text{comp} r u u_0} \mapsto Q_{hc} \\
\end{array} \right\} : C | \Theta \leadsto \Sigma_{n+1} \]

Vezzosi et al. “Cubical Agda: A Dependently Typed Programming Language with Univalence and Higher Inductive Types” ICFP ’19
The Future
Change Is Happening

• Software systems tend to be big, slow, and buggy

• There are broad forces at work that are compelling changes
  • Security and the increasing dependence on software
  • Revolution in the underlying hardware
  • A perpetual shortage of skilled programmers
Change Is Coming

• Software systems are big, slow, and buggy
  • And getting more so

• There are broad forces at work that are compelling changes
  • Security and the increasing dependence on software
    • Software verification
  • Revolution in the underlying hardware
    • Moving beyond Turing languages
  • A perpetual shortage of skilled programmers
    • Automation of programming
This Course

• Convey ethos of programming languages as a topic of study
  • And some of the important techniques

• Illustrate these ideas with examples
  • From current practice and research

• See the future before it happens
Modular design is the key to successful programming. When writing a modular program to solve a problem, one first divides the problem into subproblems, then solves the subproblems, and finally combines the solutions.

The ways in which one can divide up the original problem depend directly on the ways in which one can glue solutions together. Therefore, to increase one’s ability to modularize a problem conceptually, one must provide new kinds of glue in the programming language.

fun innerproduct(a, b, n):
    c := 0
    for i := 1 step 1 until n do
        c := c + a[i] * b[i]
    return c

• Statements operate on invisible state
• Computes word-at-a-time by repetition of assignment/modification
• Requires names for arguments, iterator, return value

let innerproduct = zip |> (map *) |> (reduce +)

• Built from composable functions (map, reduce, pipe)
• Operates on whole conceptual units (lists), no repeated steps
• No names for arguments or temporaries
Radically Different Hardware is Here
New Kinds of Glue are Coming

- Java → Scala + K
- Objective-C → Swift
- C++ → R
- JS → TS
- R → Julia
- Python → Python (with types)
More Reliable Software is Coming

Messenger used to receive bugs reports on a daily basis; since the introduction of Reason, there have been a total of 10 bugs (that's during the whole year, not per week)! Refactoring speed went from days to hours to dozens of minutes.

— “Messenger.com Now 50% Converted to Reason” 2017

Being able to encode constraints of your application in the type system makes it possible to refactor, modify, or replace large swaths of code with confidence. Rust's error model forces developers to handle every corner case. [Our system] needs very little attention. We were able to leave it running without any issues through the holiday break.

— “Rust at OneSignal” 2017
Performance of various verified symmetric crypto / hash implementations

Fastest OpenSSL assembly code
The world of software will change significantly

The changes are driven by
- New ideas in programming
- Changes in underlying hardware
- Changes in needs (e.g., security)

In this course we will focus on
- The new programming ideas
- And the intellectual tools to understand the next generation of ideas