Welcome!

Simone Campanoni
simonec@eecs.northwestern.edu
Welcome
Who we are

Simone Campanoni

Tommy McMichten
Outline

• Structure of the course

• Compilers

• Compiler IRs
CC in a nutshell

- **CS 322**: main blocks of modern compilers
  - Satisfy the system breadth for CS major
- **When**: Tuesday/Thursday 5pm - 6:20pm
- **Where**: Zoom
- **Office hours**:
  - Tommy:  Wednesday 10:00am - 11:00am
  - Simone: Thursday 6:30pm – 7:30pm
- CC is on Canvas
  - Materials/Assignments/Grades on Canvas
  - You’ll upload your assignments on Canvas
- CC is part of the sequence of compiler classes at Northwestern University
  - Other compiler-heavy classes: CS 323 and CS 397/497
  - My teaching philosophy (e.g., learn by building): link
CC materials

- Slides
- Books
- Papers and library documentation for further information
CC slides

- You can find last year slides from the [class website](#)

- We improve slides every year
  - based on problems we observe the year before
  - So: we will ask your feedbacks at the end
  - Our goal: maximize how much you learn in 10 weeks

- We will upload to Canvas the new version of the slides after each class
The CC structure

- Today

- Topic & homework

- Tuesday: Homework tests

- Thursday: Homework

- Needs to be done within 48 hours

- Needs to be done within 6 days
Output of your work

Homework after homework

you’ll **build**

your **own** compiler

from **scratch**

---

Source code (C like)

Homework N

... 

Homework 2

Homework 1

Target code (x86_64)
Homework

Each assignment is composed by:
1. A set of programs written in the source programming language (PL) considered
   (program assignment)

2. A compiler that translates the source PL to the destination PL (compiler assignment)
Homework

• Program assignment (when I’ll mention in the class)
  • You need to write Y programs
    in the source language of that assignment
  Deadline: 2 days

• Compiler assignment
  • Day X: you have the assignment
  • Deadline: 6 days after
  • Your compiler has to pass all tests included in the framework

• Late submission: you cannot be selected as a panelist
Evaluation of your work

For each assignment, you get 1 point iff:
1. Your tests are correct
2. You pass all tests using your current and prior work and
3. I will not find a bug in your implementation (I will manually inspect your code)

Some assignments can be passed either:
- **Properly**: by implementing the algorithm discussed in class
- **Naively**: you will not get the point, but you can access the next assignment (*do not submit naïve solutions*)
The CC competition

• At the end, there will be a competition between your compilers

• The team that designed the best compiler
  • Get an A automatically (no matter how many points they have)
  • Their names go to the “hall of fame” of this class
The CC grading

- 8 assignments (8 points)
  - If not submitted on time, you cannot be selected for being a panelist
- +1 point if you submit the last assignment on time for the final competition
- 4 panelist experiences (4 points)

1. Manager
2. Manager supports
3. Secretary

<table>
<thead>
<tr>
<th>Grade</th>
<th>Passed</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>&gt;= 12</td>
</tr>
<tr>
<td>A -</td>
<td>10 - 11</td>
</tr>
<tr>
<td>B +</td>
<td>8 - 9</td>
</tr>
<tr>
<td>B</td>
<td>7</td>
</tr>
<tr>
<td>C</td>
<td>6</td>
</tr>
<tr>
<td>D</td>
<td>5</td>
</tr>
<tr>
<td>F</td>
<td>0 – 4</td>
</tr>
</tbody>
</table>
Rules for homework

• You are encouraged (but not required) to work in pairs
  • Pair programming is not team programming
  • Declare your pair by the next lecture (send message via Slack to TA)
    • After this deadline, you can only split (no new/merging pairs is allowed)
    • If you don’t declare your pair, then you’ll work alone

• No copying of code is allowed between pairs

• Tool, infrastructure help is allowed between pairs
  • First try it on your own
    (google and tool documentation are your friends)

• Avoid plagiarism
  www.northwestern.edu/provost/policies/academic-integrity/how-to-avoid-plagiarism.html

• If you don’t know, please ask
Summary

• My duties
  • Teach you the blocks of a compiler
  • And how to implement them

• Your duties
  • Learn all compiler blocks presented in class
  • Implement a few of them (the most important ones)
    • Write code in C++
    • Test your code
    • Then, think much harder about how to actually test your code
    • Be ready for being in a panel when asked (the day before)
Structure & flexibility

• CC is structured w/ topics

• Best way to learn is to be excited about a topic

• Interested in something?

  **Speak**

  I’ll do my best to include your topic on the fly
Week 1

Today
- Structure
- Intro to compilers
- L1

Thursday
- Compiler structure
- Parsing
- From L1 to x86_64
Outline

• Structure of the course

• Compilers

• Compiler IRs
The role of compilers
The role of compilers

If there is no coffee, if I still have work to do, I’ll keep working, I’ll go to the coffee shop

If there is no coffee{
if I still have work to do{
  I’ll keep working;
}
I’ll go to the coffee shop;
If there is no coffee{
  if I still have work to do{
    I’ll keep working;
  } else{
    I’ll go to the coffee shop;
  }
}

Interpreter

👍 Relatively simple to build and maintain
👍 Great for prototyping
👎 It quickly becomes too slow
👎 It often consumes a lot of memory

While (s = nextStatement(){
  switch (s.type){
    case IF: ...
    case ADDITION: ...
    case RETURN: ...
  }
})
The role of compilers

If there is no coffee{
  if I still have work to do{
    I’ll keep working;
  }
  I’ll go to the coffee shop;
}

👍 Great performance of the generated binaries that get optimized
👍 Important energy savings are unlocked for the generated binaries
👎 Compilation/optimization time can increase significantly when large programs are compiled
👎 Compilers are large and complex codebases
Compiler goals

• Goal #1: correctness
• Goal #2: maximize performance and/or energy consumptions
• Goal #3: easy to be extended to
  • New architecture features (e.g., x86_64, +AVX, +TSX)
  • Evolutions of the targeted PL (e.g., C++99, C++11, C++14, C++17)
  • New architecture / ISA (e.g., RISC V)
  • New PL (e.g., Rust, Swift)
• Goal #4: Minimize maintainability costs
  • Write DRY code (Don’t Repeat Yourself)
  • Exploit code generation
Goals of your compilers in this class

• Goal #1: correctness
• Goal #2: maximize performance and/or energy consumptions
• Goal #3: easy to be extended to
  • New architecture features (e.g., x86_64, +AVX, +TSX)
  • Evolutions of the targeted PL (e.g., C++99, C++11, C++14, C++17)
  • New architecture / ISA (e.g., RISC V)
  • New PL (e.g., Rust, Swift)
• Goal #4: Minimize maintainability costs
  • Write DRY code (Don’t Repeat Yourself)
  • Exploit code generation
Structure of a compiler

Character stream (Source code)

Lexical analysis

Tokens

Syntactic & semantic analysis

AST

```
int main () {
    printf("Hello World!\n");
    return 0;
}
```
Structure of a compiler

Character stream (Source code)

Lexical analysis

Tokens

Syntactic & semantic analysis

AST

int main ...

int space string space ...

Function signature

Return type

Function name

INT

STRING
Structure of a compiler

Syntactic & semantic analysis

AST

IR code generation

Function signature

Return type

Function name

myVarX = 40
myVarY = myVarX + 2

IR
Structure of a compiler

Character stream (Source code)

Front-end

IR

Middle-end

IR

Back-end

Machine code

**EECS 322: Compiler Construction**

myVarX = 40
myVarY = myVarX + 2

**EECS 323: Code analysis and transformation**

myVarY = 42

**EECS 322: Compiler Construction**

010101110101010101
Outline

• Structure of the course

• Compilers

• Compiler IRs
define i64 @f (i64 %p0) {
    entry:
        %myVar1 = add i64 %p0, 1
    ret i64 %myVar1
}
Another example of IR

define int64 :f (int64 %p0) {
    :entry
    int64 %myVar1
    %myVar1 <- %p0 + 1
    return %myVar1
}

Multiple IRs used together

Programming language

Translation

IR1

Translation

IR2

Translation

Machine code
IRs are languages

Source code

Translation 0

...

Translation N - 1

L1

Translation N

Target code

• A compiler is a sequence of passes

• Each pass translates from a source language to a target language

• Source and target languages can be the same (transformations in the middle end)

• Some languages have the support to be written/read into/from files
In this class

- The final compiler is built as a sequence of internal compilers
- Each internal compiler translates from a source language to a target language
- Source and target languages are always different
- All languages are written/read into/from files
- Each homework is a standalone compiler
In this class

- All compilers you will build can assume the program given as input is correct
  - No need to check program’s correctness
  - Production compilers first check program’s correctness, then they do the translations/optimizations/code generations

- When you write a program in a given language, it is your job to guarantee the correctness of the program you have written

- When a compiler generates the code in its target language, it is its responsibility to generate correct code in the target language (while assuming the correctness of the code written in the source language given as input)
Let’s build our first compiler
The recipe of a disaster

1. Let’s translate independently a statement of the source program to a sequence of IR instructions

2. Let’s translate independently an IR instruction to a sequence of machine code instructions
The **good** and the **bad** compiler

```c
int main (int argc, char *argv[])
{
    return argc + 1;
}
```

### Naïve compiler

```asm
push %rbp
mov %rsp,%rbp
movl $0x0,-0x4(%rbp)
mov %edi,-0x8(%rbp)
mov %rsi,-0x10(%rbp)
lea 0x1(%rdi), %eax
add $0x1,%edi
mov %edi,%eax
pop %rbp
retq
```

### clang

```asm
lea 0x1(%rdi), %eax
retq
```

- Would you use a new PL if the resulting code is 100x slower compared to a C++ version?
- Would you use a CPU if your code is 100x slower compared to running it on an Intel CPU?
Conclusion

• Compilers translate a source language to a destination language
  • Front-end -> IR -> Middle-end -> IR -> back-end

• They help developers to be productive
  (enabling new PLs and abstractions)

• They help systems to run faster
  (enabling new resources of new CPUs)

• Correctness, efficiency (generated code and compiler itself),
  maintainability, extensibility
  are all aspects to consider when designing a compiler