LLVM Compiler System

LLVM: A Compilation Framework for Lifelong Program Analysis & Transformation

Chris Lattner Vikram Adve
University of Illinois at Urbana-Champaign
{lattner,vadve}@cs.uiuc.edu
http://llvm.cs.uiuc.edu/

ABSTRACT

This paper describes LLVM (Low Level Virtual Machine), a compiler framework designed to support *transparent, lifelong program analysis and transformation* for arbitrary programs, by providing high-level information to compiler transformations at compile-time, link-time, run-time, and in idle time between runs. LLVM defines a common, low-level code representation in Static Single Assignment (SSA) form, with several novel features: a simple, *language-independent* type-system that exposes the primitives commonly used to implement high-level language features; an instruction for typed address arithmetic; and a simple mechanism that can be used to implement the exception handling features of high-level languages (and `setjmp/longjmp` in C) uniformly and efficiently. The LLVM compiler framework and code representation together provide a combination of key capabilities that are important for practical, lifelong analysis and transformation of programs. To our knowledge, no existing compilation approach provides all these capabilities. We describe the design of the LLVM representation and compiler

mizations performed at link-time (to preserve the benefits of separate compilation), machine-dependent optimizations at install time on each system, dynamic optimization at runtime, and profile-guided optimization between runs ("idle time") using profile information collected from the end-user.

Program optimization is not the only use for lifelong analysis and transformation. Other applications of static analysis are fundamentally interprocedural, and are therefore most convenient to perform at link-time (examples include static debugging, static leak detection [24], and memory management transformations [30]). Sophisticated analyses and transformations are being developed to enforce program safety, but must be done at software installation time or load-time [19]. Allowing lifelong reoptimization of the program gives architects the power to evolve processors and exposed interfaces in more flexible ways [11, 20], while allowing legacy applications to run *well* on new systems.

This paper presents **LLVM** — Low-Level Virtual Machine — a compiler framework that aims to make lifelong program analysis and transformation available for arbitrary

Low Level Virtual Machine

- Started around 2000 at University of Illinois at Urbana-Champaign by Chris Lattner, first public release around 2003
- Now maintained by Apple
  - Default compiler for OS X and IOS
- Users & contributors include: Sony, Adobe, Intel, NVIDIA, XMOS, and many others
- Used in various scenarios – including both AOT and JIT compilers, code analysis, etc.

Low Level Virtual Machine

- …d 2000 at University of Illinois at Urbana-
  … Chris Lattner, first public release
- …ined by Apple
  - …ompiler for OS X and IOS
- …ributors include: Sony, Adobe, Intel,
  …S, and many others
- Used in various scenarios – including both AOT and JIT compilers, code analysis, etc.

Low **L**evel **V**irtual **M**achine

- Started around 2000 at University of Illinois at Urbana-Champaign by Chris Lattner, first public release around 2003
- Now maintained by Apple
  - Default compiler for OS X and IOS
- Users & contributors include: Sony, Adobe, Intel, NVIDIA, XMOS, and many others
- Used in various scenarios – including both AOT and JIT compilers, code analysis, etc.

Low **L**evel **V**irtual **M**achine

- Permissive license (BSD-style)
- Modern compiler, written in C++

- Under active development

- Modular design
- Well documented
- Language agnostic
  - Does not care about the frontend
  - Many targets available (x86, ARM, PowerPC, etc.)

Low Level Virtual Machine

- Permissive license (BSD-style)
- Modern compiler, written in C++
- Under active development
- Modular design
- Well documented
- Language agnostic
  - Does not care about the frontend
  - Many targets available (x86, ARM, PowerPC, etc.)

- Templates & advanced C++ constructs
- Indecipherable error messages
- On a plus side, does not use BOOST

Low **L**evel **V**irtual **M**achine

- Permissive license (BSD-style)
- Modern compiler, written in C++
  - Templates & advanced C++ constructs
  - Indecipherable error messages
  - On a plus side, does not use BOOST
- Under active development
  - Backwards compatibility is frowned upon
- Modular design
- Well documented
- Language agnostic
  - Does not care about the frontend
  - Many targets available (x86, ARM, PowerPC, etc.)

Low Level Virtual Machine

- Permissive license (BSD-style)
- Modern compiler, written in C++
  - Templates & advanced C++ constructs
  - Indecipherable error messages
  - On a plus side, does not use BOOST
- Under active development
  - Backwards compatibility is frowned upon
- Modular design - kind of
- Well documented
- Language agnostic
  - Does not care about the frontend
  - Many targets available (x86, ARM, PowerPC, etc.)

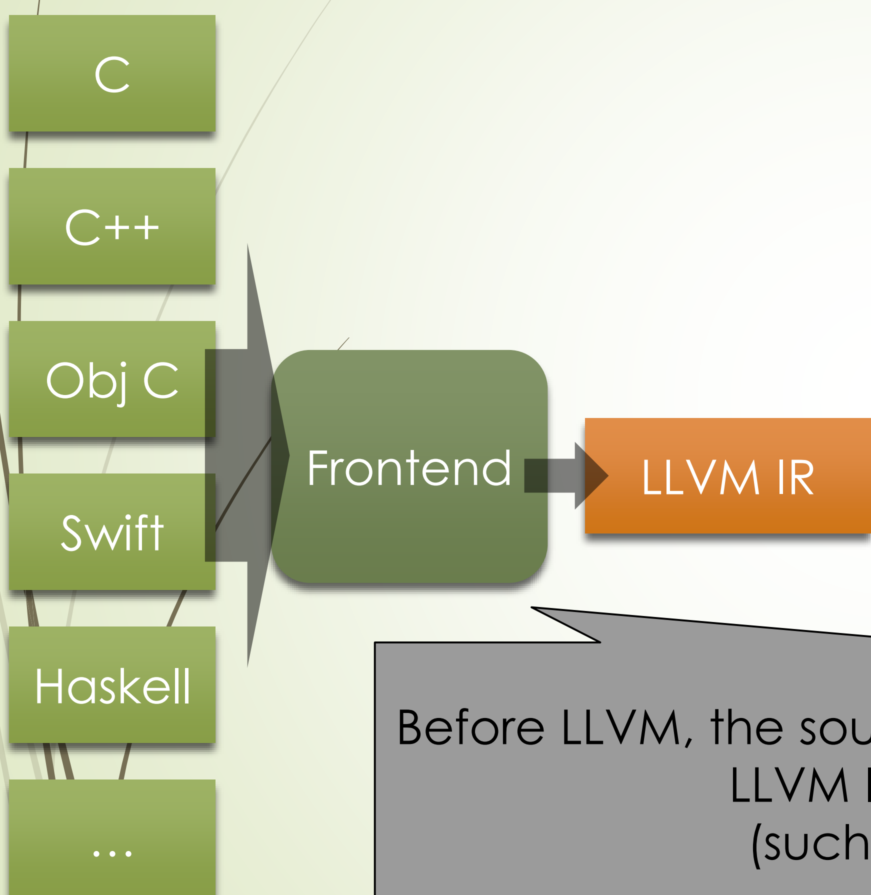
Low Level Virtual Machine

- Permissive license (BSD-style)
- Modern compiler, written in C++
  - Templates & advanced C++ constructs
  - Indecipherable error messages
  - On a plus side, does not use BOOST
- Under active development
  - Backwards compatibility is frowned upon
- Modular design - kind of
- ~~Well~~ documented – just better than others, but the bar is low
- Language agnostic
  - Does not care about the frontend
  - Many targets available (x86, ARM, PowerPC, etc.)

Low **L**evel **V**irtual **M**achine

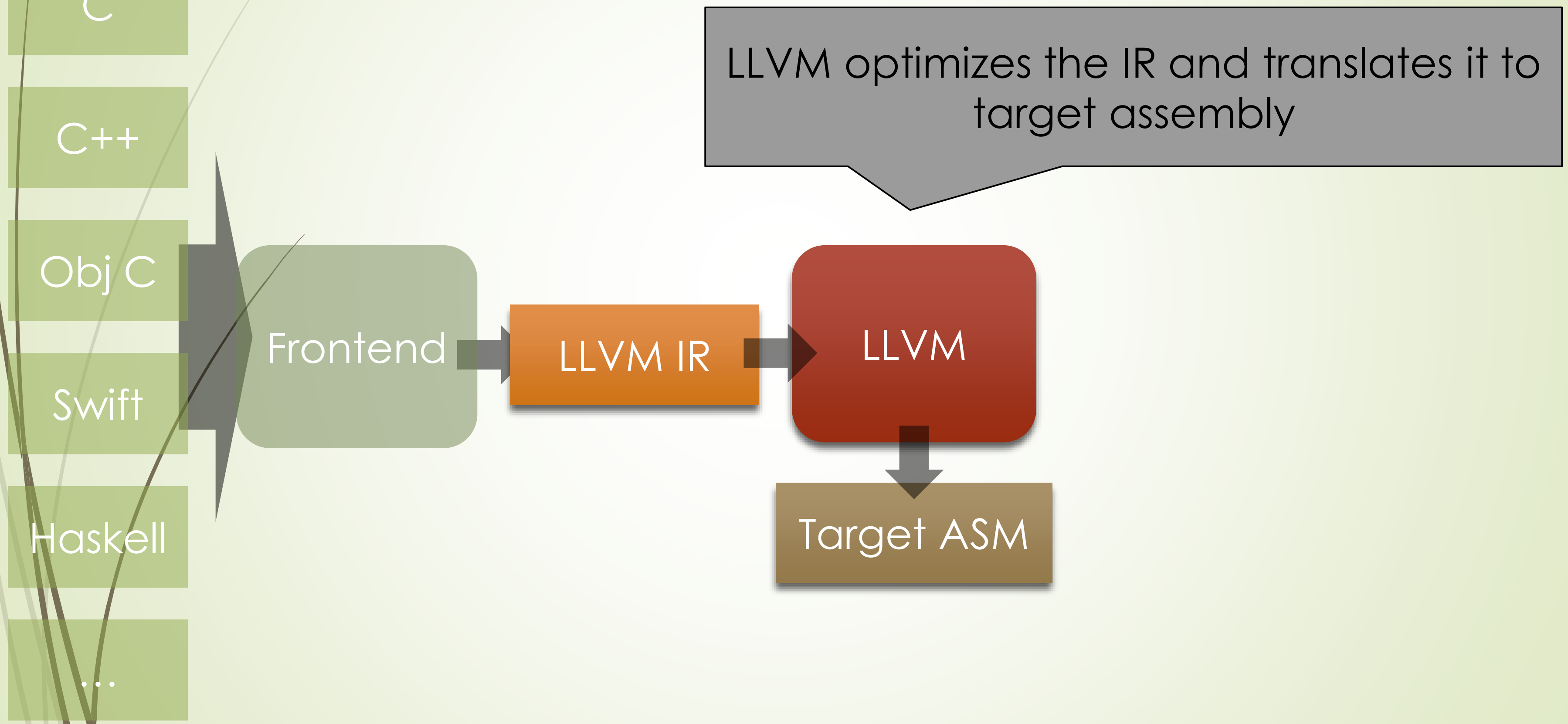
- C
- C++
- Obj C
- Swift
- Haskell
- …

Low **L**evel **V**irtual **M**achine

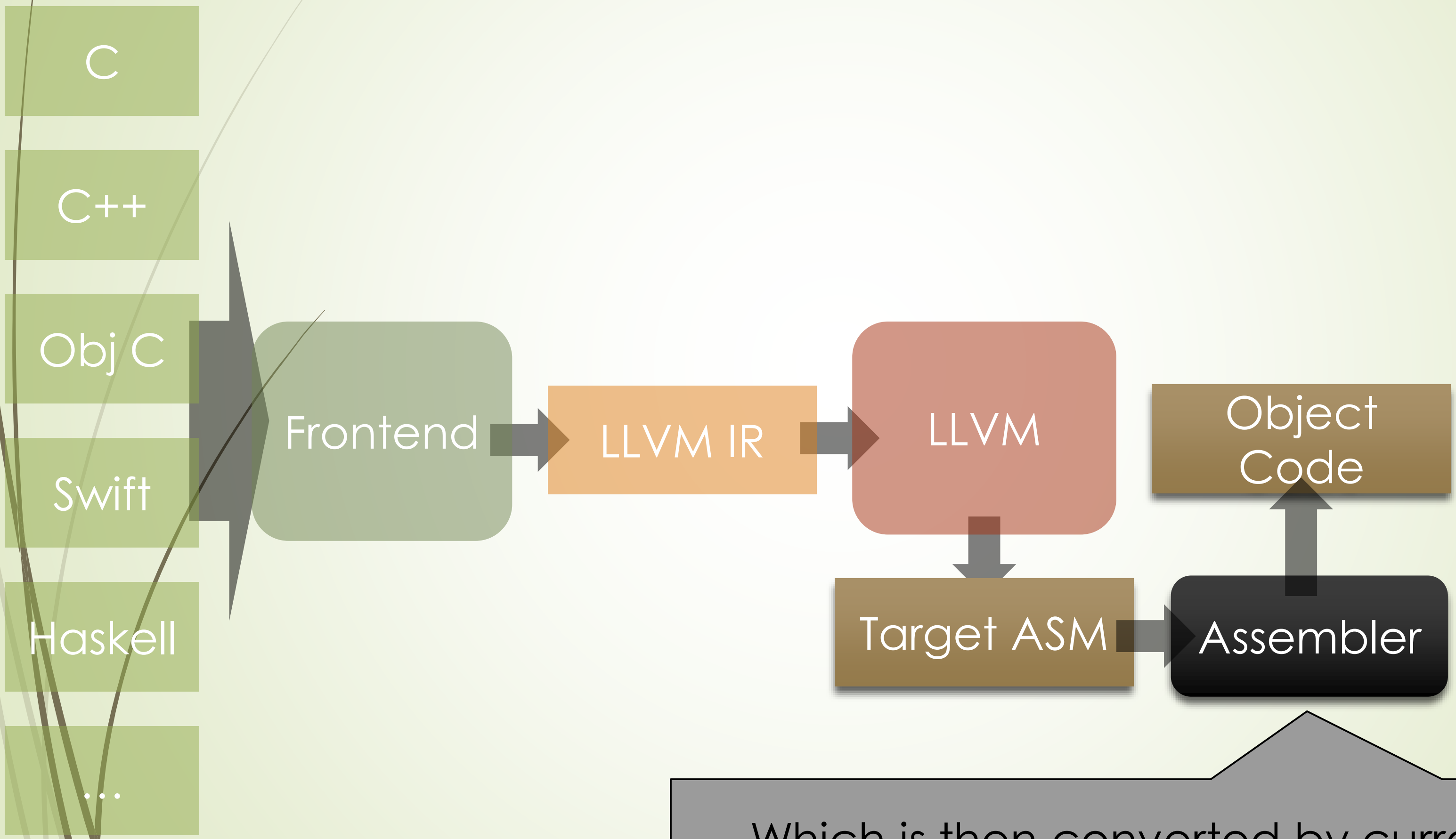


Before LLVM, the source language must be translated to LLVM IR by custom frontend (such as Clang for C/C++)

Low **L**evel **V**irtual **M**achine

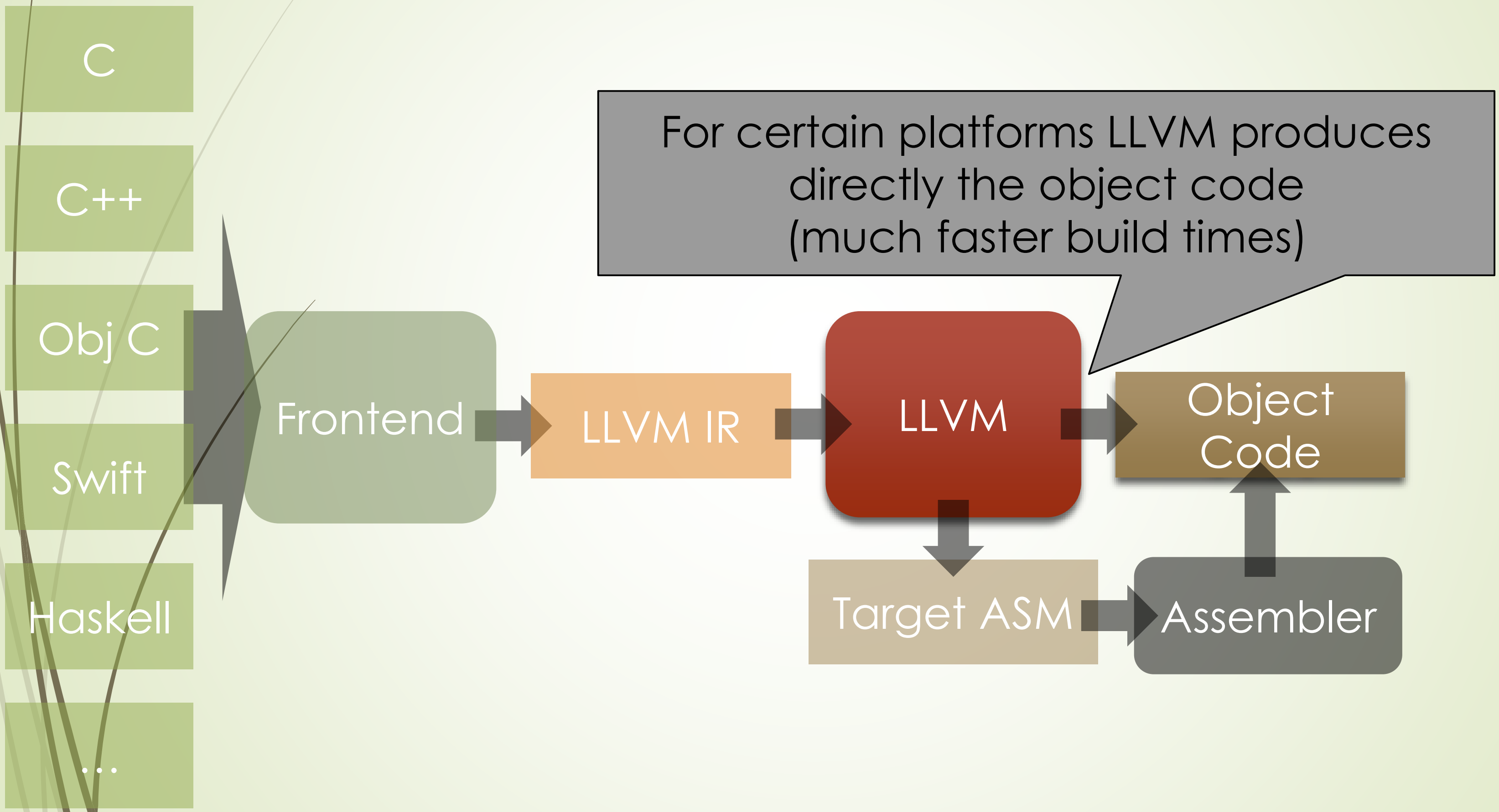


Low **L**evel **V**irtual **M**achine

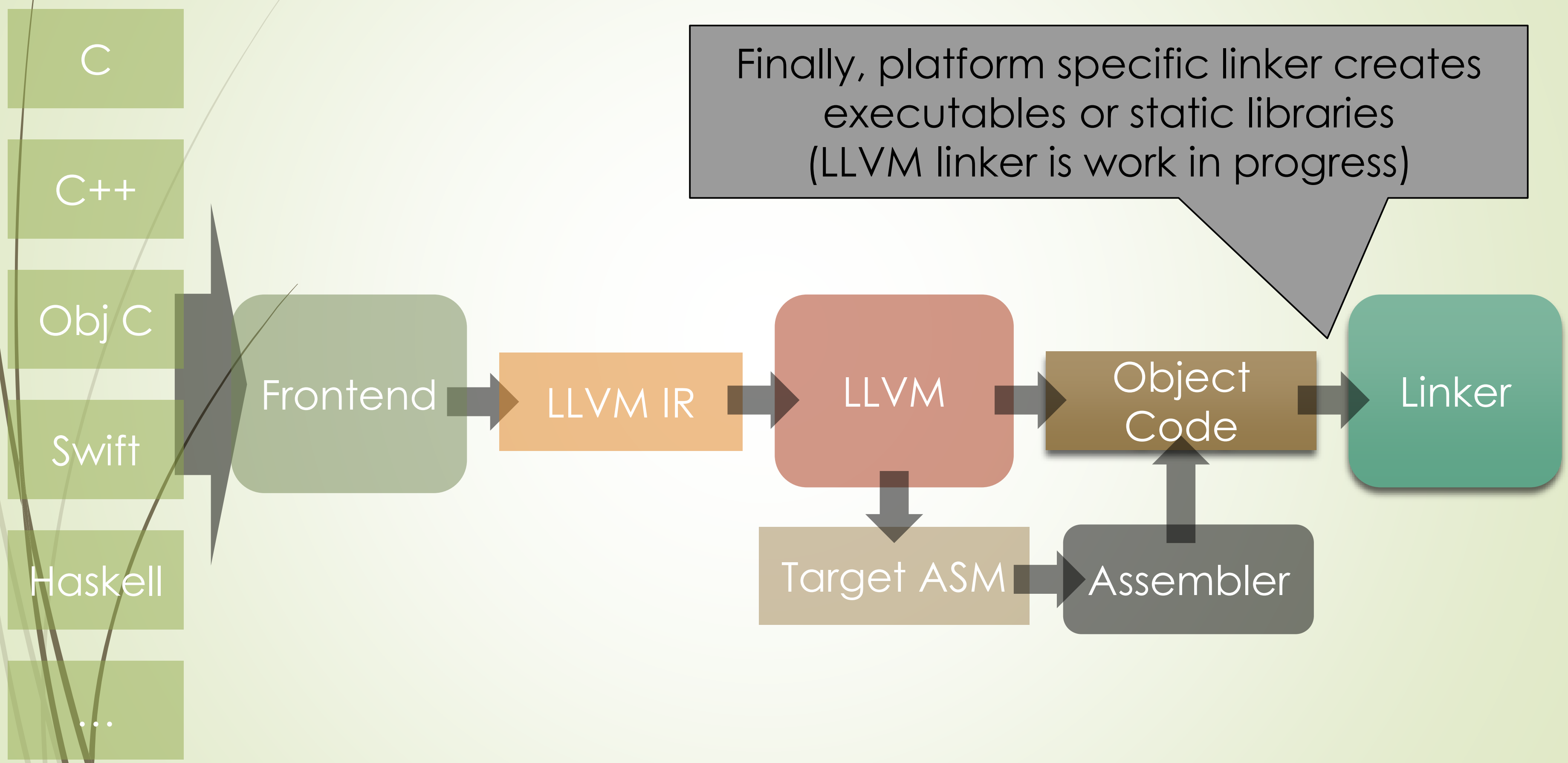


Which is then converted by current platform assembler to object code

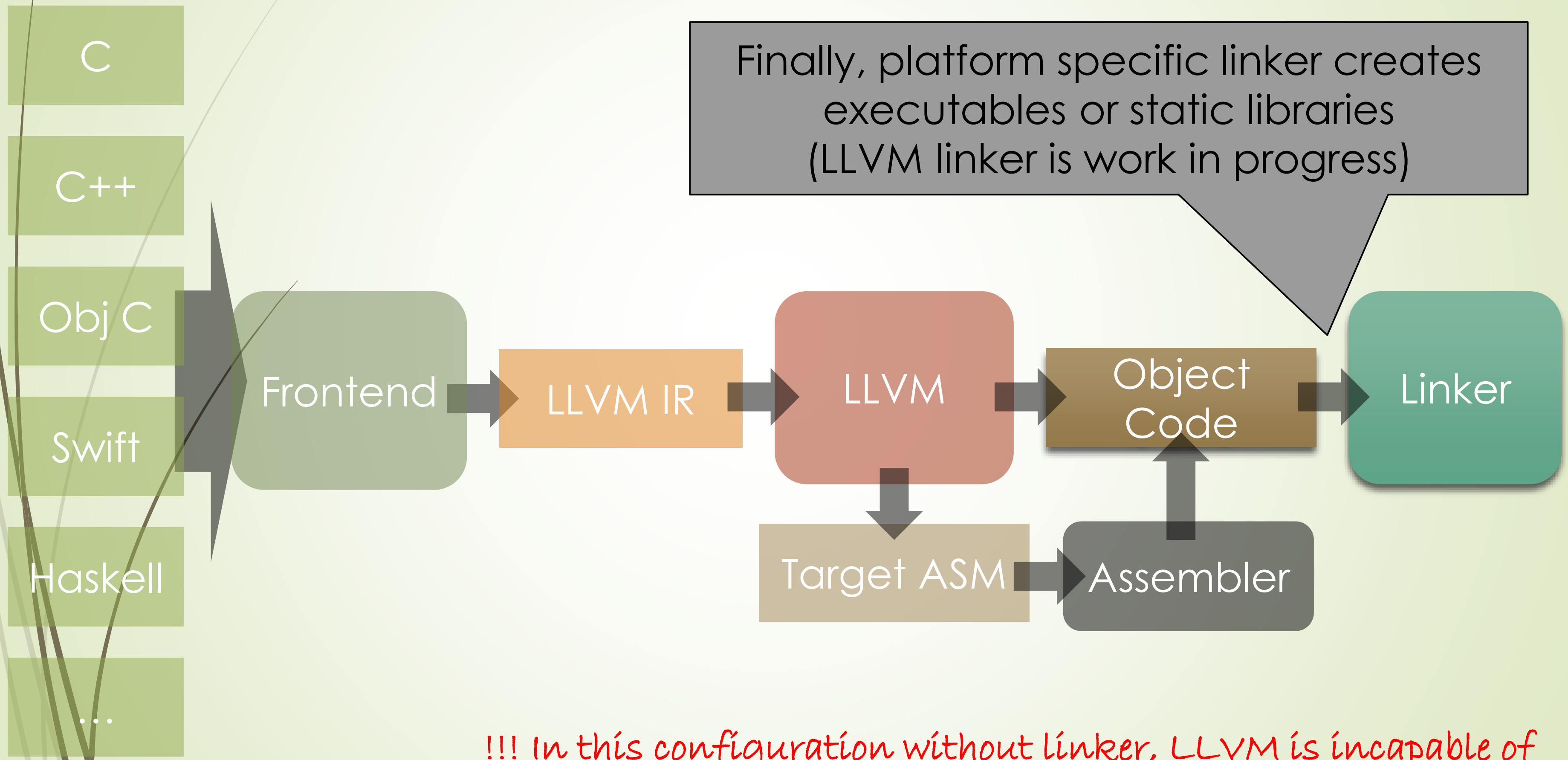
Low Level Virtual Machine



Low Level Virtual Machine



Low **L**evel **V**irtual **M**achine



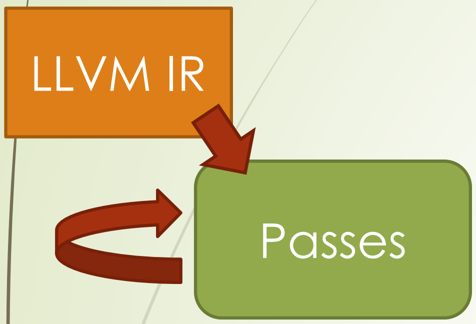
!!! In this configuration without linker, LLVM is incapable of link time optimizations !!!

Low Level Virtual Machine

- Does not contain frontends
- Or linker (to be changed soon)
- Or assemblers in many platforms
  - Using platform assembler instead

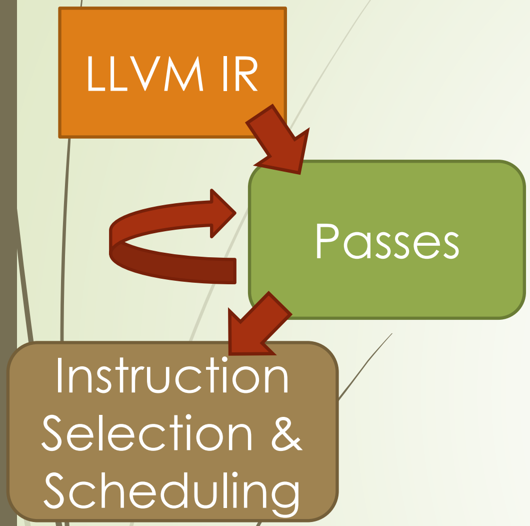
- LLVM's main focus is on IR optimizations and backend up to assembly
  - SSA IR optimizations
  - Instruction scheduling
  - Register allocation
  - Low level peephole optimizations (to some extent)

LLVM workflow



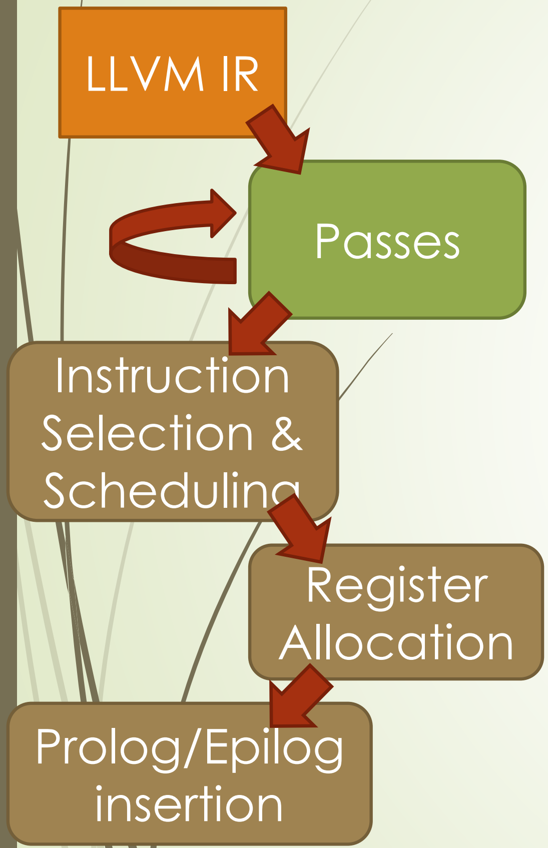
- Several analysis (readonly) and optimization passes are performed over the IR
- Passes can run on
  - Modules
  - Functions
  - Basic blocks
- Passes can
  - Depend on other passes
  - Preserve other passes
- LLVM provides a scheduler that calls the passes in a most effective way

LLVM workflow



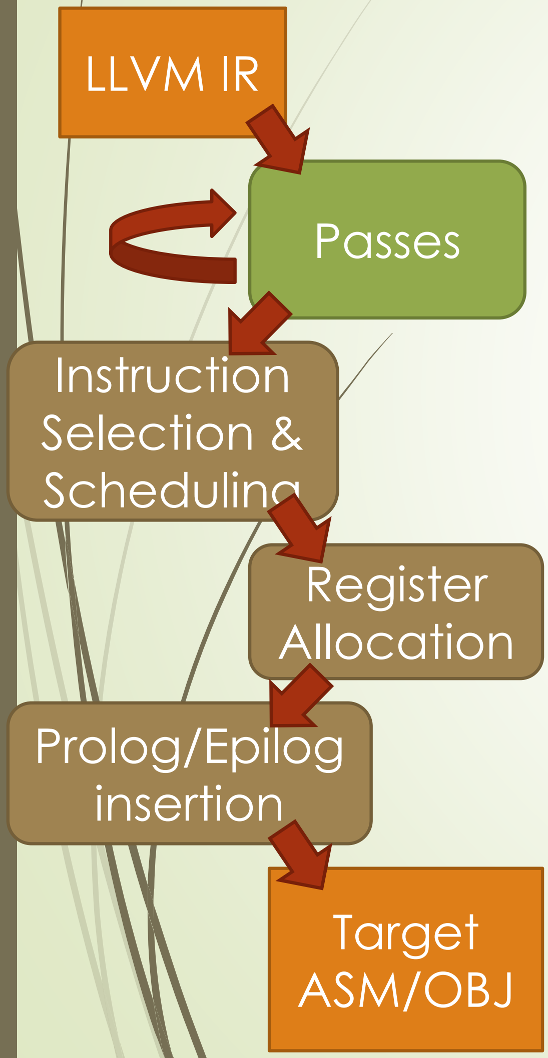
- When IR passes are finished, LLVM enters the target specific backend
- Target specific instructions are selected and attached to the IR
- Afterwards, target specific SSA passes may execute

LLVM workflow



- After SSA target passes, register allocator is executed
  - Linear scan
  - Greedy
  - Region-based
  - GRA
- Function Prologues and epilogues are inserted
- After these steps, late target optimizations are usually executed
  - peepholer

LLVM workflow



- Finally IR is dumped as target object, or assembly files

Bitcode Architecture

Bitcode organization

- Code in LLVM is organized into modules (akin to compilation units)
- Each module contains global variables and functions
  - Unique names
  - Linkage specification
  - Calling convention
  - Visibility,
  - Etc.
- Inside functions, code is organized into basic blocks

Basic Blocks

- Consecutive sequence of instructions with only one entry and one exit point
- Entry point
  - At the beginning, the only instruction allowed to be jump target
- Exit point
  - At the end, the only instruction allowed to be jump (or return, throw, etc.)
- Once basic block is entered, all its instructions are guaranteed to be executed

Basic Blocks

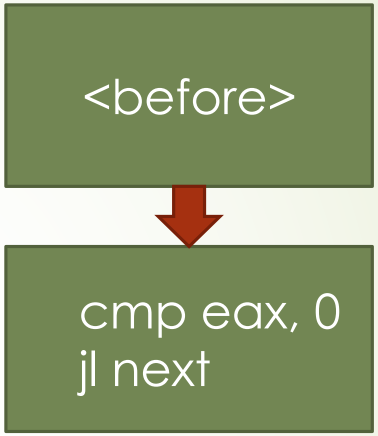
```
    <before>
cond:
    cmp eax, 0
    jl next
    <body>
    jmp cond
next:
    <after>
```

Basic Blocks

```
    <before>
cond:
    cmp eax, 0
    jl next
    <body>
    jmp cond
next:
    <after>
```

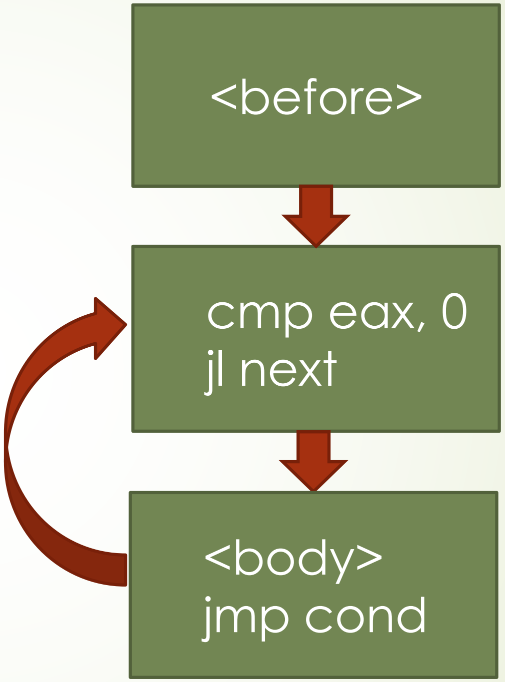
Basic Blocks

```
    <before>
cond:
    cmp eax, 0
    jl next
    <body>
    jmp cond
next:
    <after>
```



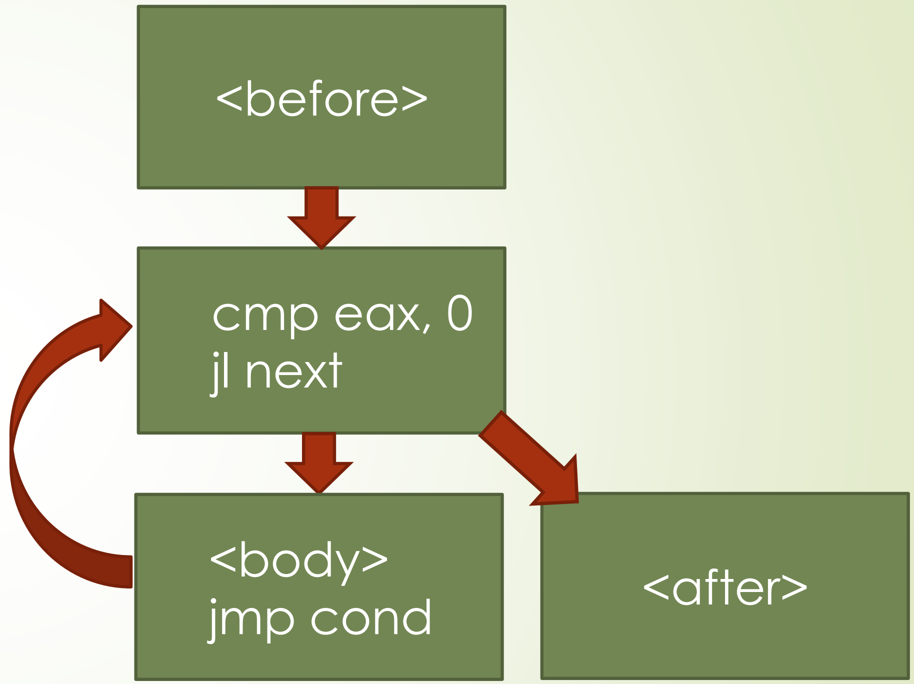
Basic Blocks

```
    <before>
cond:
    cmp eax, 0
    jl next
    <body>
    jmp cond
next:
    <after>
```



Basic Blocks

```
    <before>
cond:
    cmp eax, 0
    jl next
    <body>
    jmp cond
next:
    <after>
```



Types

- void (no value, no size)
- integer (variable size in bits)
  - i1, i32, i67837
  - integers do not care about their sign
- Floats
  - half, float, double, …
- Arrays (of same element types, fixed size)
  - [10 x i32], [10 x [10 x i8]]
- Structures
  - { i32, i32, float, i1 }
  - <{ i8, i8, i32 }> packed structure (padding=0, align=1byte)
- Functions
  - i32 (i32, i32)
- Pointers (*)

Types

+ fp128
+ fp80 (x86)
+ ppc 128bit fp
+ mmx types
+ vectors (SIMD)
+ labels, tokens, metadata, ...

- void (no value, no size)
- integer (variable size in bits)
  - i1, i32, i67837
  - integers do not care about their sign
- Floats
  - half, float, double, …
- Arrays (of same element types, fixed size)
  - [10 x i32], [10 x [10 x i8]]
- Structures
  - { i32, i32, float, i1 }
  - <{ i8, i8, i32 }> packed structure (padding=0, align=1byte)
- Functions
  - i32 (i32, i32)
- Pointers (*)

Variables

- LLVM provides three types of variables
- Stack allocated variables
  - `alloca` = creates variable on stack and returns a pointer to it
  - `load` & `store`
- Global variables (prefixed with @)
  - Contain pointers to the global variables (i.e. similar to stack allocated variables)
- Local variables (prefixed with %)
  - Result of each instruction goes into new local variable
  - In fact, the variable and the instruction are the same thing in LLVM
  - This makes the LLVM variables SSA values

Single Static Assignment

- Each variable is defined exactly once
- Each variable's definition dominates all its uses

(Cytron, Ferrante, Rosen, Wegman, Zadeck, IBM & Brown 1989)

Efficiently Computing Static Single Assignment Form and the Control Dependence Graph

RON CYTRON, JEANNE FERRANTE, BARRY K. ROSEN, and MARK N. WEGMAN
IBM Research Division
and
F. KENNETH ZADECK
Brown University

In optimizing compilers, data structure choices directly influence the power and efficiency of practical program optimization. A poor choice of data structure can inhibit optimization or slow compilation to the point that advanced optimization features become undesirable. Recently, static single assignment form and the control dependence graph have been proposed to represent

Single Static Assignment

- Each variable is defined exactly once
- Each variable's definition dominates all its uses

  (Cytron, Ferrante, Rosen, Wegman, Zadeck, IBM & Brown 1989)

- Over its lifetime, a variable is replaced by its versions, new version created whenever the variable is assigned:

```
x = 1                  x1 = 1
x = 2        ->        x2 = 2
z = x                  z1 = x2
```

Single Static Assignment

- Each variable is defined exactly once
- Each variable's definition dominates all its uses

  (Cytron, Ferrante, Rosen, Wegman, Zadeck, IBM & Brown 1989)

- Over its lifetime, a variable is replaced by its versions, new version created whenever the variable is assigned:

```
x = 1                  x1 = 1
x = 2       ->         x2 = 2
z = x                  z1 = x2
```

This nicely simplifies the use-def chains (so we can trivially see that x1 is never used and thus eliminate the assignment

PHI nodes

- SSA is simple when we have only sequential code, but what if a value may come from different places:

```
if (…)
  a = 2;
else
  a = 3;
b = a
```

PHI nodes

- SSA is simple when we have only sequential code, but what if a value may come from different places:

```
if (…)
  a = 2;
else
  a = 3;
b = a
```

->

```
if (…)
  a1 = 2;
else
  a2 = 3;
b1 = ???
```

PHI nodes

- SSA is simple when we have only sequential code, but what if a value may come from different places:

```
if (…)                      if (…)
  a = 2;                      a1 = 2;
else              ->        else
  a = 3;                      a2 = 3;
b = a                       b1 = ???
```

- PHI nodes generate new values by selecting “by magic” the proper values based on where the control flow arrived from

PHI nodes

- SSA is simple when we have only sequential code, but what if a value may come from different places:

```
if (…)                    if (…)
  a = 2;                    a1 = 2;
else           ->         else
  a = 3;                    a2 = 3;
b = a                     a3 = phi(a1,a2)
                          b1 = a3
```

- PHI nodes generate new values by selecting “by magic” the proper values based on where the control flow arrived from

Bitcode Instructions

- Terminator instructions
  - Terminate basic block (jumps, returns, exceptions)

Bitcode Instructions

- Terminator instructions
  - Terminate basic block (jumps, returns, exceptions)
- Binary Operations
  - Different opcodes for floats (fadd,…) and integers (add…)

Bitcode Instructions

- Terminator instructions
  - Terminate basic block (jumps, returns, exceptions)
- Binary Operations
  - Different opcodes for floats (fadd,…) and integers (add…)
- Shifts, rotations
  - arithmetic, logic

Bitcode Instructions

- Terminator instructions
  - Terminate basic block (jumps, returns, exceptions)
- Binary Operations
  - Different opcodes for floats (fadd,…) and integers (add…)
- Shifts, rotations
  - arithmetic, logic
- Vector operations
  - Extract, insert, shuffle

Bitcode Instructions

- Terminator instructions
  - Terminate basic block (jumps, returns, exceptions)
- Binary Operations
  - Different opcodes for floats (fadd,…) and integers (add…)
- Shifts, rotations
  - arithmetic, logic
- Vector operations
  - Extract, insert, shuffle
- Memory access & addressing
  - allocation, load, store, fences, getelementptr

Bitcode Instructions

- Terminator instructions
  - Terminate basic block (jumps, returns, exceptions)
- Binary Operations
  - Different opcodes for floats (fadd,…) and integers (add…)
- Shifts, rotations
  - arithmetic, logic
- Vector operations
  - Extract, insert, shuffle
- Memory access & addressing
  - allocation, load, store, fences, getelementptr
- Conversions
  - Zext, sext, …

Bitcode Instructions

- Terminator instructions
  - Terminate basic block (jumps, returns, exceptions)
- Binary Operations
  - Different opcodes for floats (fadd,…) and integers (add…)
- Shifts, rotations
  - arithmetic, logic
- Vector operations
  - Extract, insert, shuffle
- Memory access & addressing
  - allocation, load, store, fences, getelementptr
- Conversions
  - Zext, sext, …
- Calling

Bitcode Instructions

- Terminator instructions
  - Terminate basic block (jumps, returns, exceptions)
- Binary Operations
  - Different opcodes for floats (fadd,…) and integers (add…)
- Shifts, rotations
  - arithmetic, logic
- Vector operations
  - Extract, insert, shuffle
- Memory access & addressing
  - allocation, load, store, fences, getelementptr
- Conversions
  - Zext, sext, …
- Calling
- Comparisons
  - Icmp, fcmp, …

Bitcode Instructions

- Terminator instructions
  - Terminate basic block (jumps, returns, exceptions)
- Binary Operations
  - Different opcodes for floats (fadd,…) and integers (add…)
- Shifts, rotations
  - arithmetic, logic
- Vector operations
  - Extract, insert, shuffle
- Memory access & addressing
  - allocation, load, store, fences, getelementptr
- Conversions
  - Zext, sext, …
- Calling
- Comparisons
  - Icmp, fcmp, …
- Exceptions
  - Catchpads, landingpads, etc.

Bitcode Instructions

- Terminator instructions
  - Terminate basic block (jumps, returns, exceptions)
- Binary Operations
  - Different opcodes for floats (fadd,…) and integers (add…)
- Shifts, rotations
  - arithmetic, logic
- Vector operations
  - Extract, insert, shuffle
- Memory access & addressing
  - allocation, load, store, fences, getelementptr
- Conversions
  - Zext, sext, …
- Calling
- Comparisons
  - Icmp, fcmp, …
- Exceptions
  - Catchpads, landingpads, etc.
- Virtuals (PHI node)

Bitcode Instructions

- Terminator instructions
  - Terminate basic block (jumps, returns, exceptions)
- Binary Operations
  - Different opcodes for floats (fadd,…) and integers (add…)
- Shifts, rotations
  - arithmetic, logic
- Vector operations
  - Extract, insert, shuffle
- Memory access & addressing
  - allocation, load, store, fences, getelementptr
- Conversions
  - Zext, sext, …
- Calling
- Comparisons
  - Icmp, fcmp, …
- Exceptions
  - Catchpads, landingpads, etc.
- Virtuals (PHI node)
- Intrinsics (gc, padding, etc.)

Bitcode Instructions

- All instructions are strongly typed and types are almost always explicit in the IR
- Each instruction can be named, in which case the variable bearing the result of the instruction will carry the name

```
%2 = call i32 @min(i32 1,%1)
```

Bitcode Instructions

- All instructions are strongly typed and types are almost always explicit in the IR
- Each instruction can be named, in which case the variable bearing the result of the instruction will carry the name

```
%2 = call i32 @min(i32 1, %1)
```

Name of the variable holding result of the instruction (if no name is provided, llvm assigns unique number)

Bitcode Instructions

- All instructions are strongly typed and types are almost always explicit in the IR
- Each instruction can be named, in which case the variable bearing the result of the instruction will carry the name

Llvm instruction (function call)

```
%2 = call i32 @min(i32 1, %1)
```

Bitcode Instructions

- All instructions are strongly typed and types are almost always explicit in the IR
- Each instruction can be named, in which case the variable bearing the result of the instruction will carry the name

```
%2 = call i32 @min(i32 1, %1)
```

Type of the result of the instruction (explicit)

Bitcode Instructions

- All instructions are strongly typed and types are almost always explicit in the IR
- Each instruction can be named, in which case the variable bearing the result of the instruction will carry the name

```
%2 = call i32 @min(i32 1, %1)
```

Type of the literal argument

Bitcode Instructions

- All instructions are strongly typed and types are almost always explicit in the IR
- Each instruction can be named, in which case the variable bearing the result of the instruction will carry the name

Value of argument 1 (constant 1)

```
%2 = call i32 @min(i32 1, %1)
```

Value of argument 2 (variable 1)

Terminators

- `ret`
  - return from a function, may or may not return a value

    ```
    ret i32 %3
    ```
- `br`
  - Unconditional

    ```
    br label %4
    ```
  - Conditional (branches to two basic blocks based on condition)

    ```
    br i1 %2, label %3, label %4
    ```
- Also indirect branch, switch, exceptions throwing & catching

Binary Operators

- `add, sub, mul`
  - Do not care about signedness of operands

    ```
    %1 = add i32 %a, %b
    ```
- `udiv, sdiv`
  - Unsigned and signed division of integers
- `fadd, fsub, fmul, fdiv`
  - Floating point arithmetics

    ```
    %3 = fadd double %a, %b
    ```

Memory

- `alloca`
  - allocates space for given type on stack and returns a pointer to it

```
%1 = alloca i32, align 4
```

- `load`
  - Loads contents of given pointer to register

```
%2 = load i32 i32* %1, align 4
```

- `store`
  - stores to a pointer

```
store i32 %I, i32* %1, align 4
```

Memory

- `getelementptr`
  - address calculation for subelements of aggregate types (array, structure)
  - Different semantics for arrays and structs
  - Very important and often misunderstood instruction

```
%A = type { i32, double }
# type of %1 is %A*
# A*[3].double becomes
%3 = getelementptr %A, %A* %1, i32 3
%4 = getelementptr %A, %A* %1, i32 0, i32 1
```

Memory

- `getelementptr`
  - address calculation for subelements of aggregate types (array, structure)
  - Different semantics for arrays and

For arrays, index of element we want to access

Pointer to the aggregate type

isund

```
%A = typ                  double }
# type of %1 is
# A*[3].double becomes
%3 = getelementptr %A, %A* %1, i32 3
%4 = getelementptr   A, %A* %1, i32 0, i32 1
```

Type we are operating on

Memory

- `getelementptr`
  - address calculation for subelements of aggregate types (array, structure)
  - Different semantics for arrays and
  - Very important and often misund

```
}
# type          %A
# A*[3].doubl        es
%3 = getelementptr %A, %A* %1, i32 3
%4 = getelementptr %A, %A* %1, i32 0, i32 1
```

For structs, index of element in declaration

Pointer to the aggregate type

Type we are operating on

Memory

- `getelementptr`
  - address calculation for subelements of aggregate types (array, structure)
  - Different semantics for arrays and
  - Very important and often misunderstood instruc

```
                                    }
# type              %A
# A*[3].double      es
%3 = getelementptr %A, %A* %1, i32 3
%4 = getelementptr %A, %A* %1, i32 0, i32 1
```

For structs, index of element in declaration

???

Pointer to the aggregate type

Type we are operating on

Bitcode Representations

- Traditionally LLVM differentiates between three bitcode representations
  - Human readable LLVM (this is what we have seen so far)

    ```
    %1 = add i32 %a, %b
    ```

  - Binary bitcode (this is what LLVM toolchain members usually pass)

    ```
    0x0 f2 cd 0a 0b 54 5a 35 12 7e 2f
    ```

  - C++ API (this is what frontend developers use to construct the IR

    ```
    auto b = llvm::BasicBlock::Create(llvm::GetGlobalContext(), “”, f);
    auto li = new llvm::LoadInst(ptr, “”, false, b);
    llvm::ReturnInst::Create(llvm::getGlobalContext(), li, b);
    ```

C++ API

Everything is a class, something is a pointer

- Modules, Functions and BasicBlocks
  - `llvm::Module`, `llvm::Function`, `llvm::BasicBlock`
- Types, values and instructions are represented by classes
  - `llvm::Type`, `llvm::Value` and `llvm::Instruction`
- Usually pointers to the classes are expected
- Many (but not all) of the classes should not be created using constructors, but provide static `Create` methods

Types

- Each type must be specified and have corresponding class
- Types are common to all modules in llvm

```
auto tv=llvm::Type::getVoidTy(llvm::getGlobalContext())

auto ti=llvm::IntegerType::get(context, 32)

auto td=llvm::Type::getDoubleTy(context)

auto ts=llvm::StructType::create(context, “name”)
std::vector<llvm::Type *> fields = { ti, ti, td }
ts->setBody(fields, false);
```

Types

- Each type must be specified and have corresponding class
- Types are common to all modules in llvm

Often, context is required (in case of multiple llvm instances, we can always use global context)

```
auto tv=llvm::Type::getVoidTy(llvm::getGlobalContext())

auto ti=llvm::IntegerType::get(context, 32)

auto td=llvm::Type::getDoubleTy(context)

auto ts=llvm::StructType::create(context, “name”)
std::vector<llvm::Type *> fields = { ti, ti, td }
ts->setBody(fields, false);
```

Types

- Each type must be specified and have corresponding class
- Types are common to all modules in llvm

C++ void

```
auto tv=llvm::Type::getVoidTy(llvm::getGlobalContext())

auto ti=llvm::IntegerType::get(context, 32)

auto td=llvm::Type::getDoubleTy(context)

auto ts=llvm::StructType::create(context, “name”)
std::vector<llvm::Type *> fields = { ti, ti, td }
ts->setBody(fields, false);
```

Types

- Each type must be specified and have corresponding class
- Types are common to all modules in llvm

C++ int or unsigned

```
auto tv=llvm::Type::getVoidTy(llvm::getGlobalContext())

auto ti=llvm::IntegerType::get(context, 32)

auto td=llvm::Type::getDoubleTy(context)

auto ts=llvm::StructType::create(context, “name”)
std::vector<llvm::Type *> fields = { ti, ti, td }
ts->setBody(fields, false);
```

Types

- Each type must be specified and have corresponding class
- Types are common to all modules in llvm

C++ double

```
auto tv=llvm::Type::getVoidTy(llvm::getGlobalContext())

auto ti=llvm::IntegerType::get(context, 32)

auto td=llvm::Type::getDoubleTy(context)

auto ts=llvm::StructType::create(context, “name”)
std::vector<llvm::Type *> fields = { ti, ti, td }
ts->setBody(fields, false);
```

Types

- Each type must be specified and have corresponding class
- Types are common to all modules in llvm

```
auto tv=llvm::Type::getVoidTy(llvm::getGlobalContext())

auto ti=llvm::IntegerType::get(context, 32)

auto td=llvm::Type::getDoubleTy(context)

auto ts=llvm::StructType::create(context, “name”)
std::vector<llvm::Type *> fields = { ti, ti, td }
ts->setBody(fields, false);
```

struct { int a; int b; double c; }

Packed?

Creating constants

- Constants inherit from llvm::Value and can be used as arguments to instructions
- As types, constants live outside modules

```
llvm::ConstantInt::get(context, llvm::APInt(32, 1,
false));
```

```
llvm::ConstantFP::get(context, llvm::APFloat(3.14))
```

Creating constants

- Constants inherit from llvm::Value and can be used as arguments to instructions
- As types, constants live outside modules

size

```
llvm::ConstantInt::get(context, llvm::APInt(32, 1, false))
```

value

signed?

```
llvm::ConstantFP::get(context, llvm::APFloat(3.14))
```

Creating a function

```
int min(int i, int j) {
    return i < j ? i : j;
}
```

Creating a function

```
auto m = llvm::Module::Create(“name”, context);

auto ft = llvm::FunctionType::get(ti, { ti, ti },
false);

auto f = llvm::Function::Create(ft,
llvm::GlobalValue::ExternalLinkage, “min”, m);

f->setCallingConvention(llvm::CallingConv::C);
```

Creating a function

First create a module with given name

```
auto m = llvm::Module::Create(“name”, context);

auto ft = llvm::FunctionType::get(ti, { ti, ti },
false);

auto f = llvm::Function::Create(ft,
llvm::GlobalValue::ExternalLinkage, “min”, m);

f->setCallingConvention(llvm::CallingConv::C);
```

Creating a function

The function must have a type, in our case
int (*ptr)(int, int)

```
text);

auto ft = llvm::FunctionType::get(ti, { ti, ti },
false);

auto f = llvm::Function::Create(ft,
llvm::GlobalValue::ExternalLinkage, “min”, m);

f->setCallingConvention(llvm::CallingConv::C);
```

varargs?

Creating a function

```
auto m = llvm::Module::Create(“name”, context);
pe::get ,

auto f = llvm::Function::Create(ft,
llvm::GlobalValue::ExternalLinkage, “min”, m);
f->setCallingConvention(llvm::CallingConv::C);
```

Create the function object

type

symbol visibility

name of the function (unique)

Module the function belongs to

Creating a function

```
auto m = llvm::Module::Create("name", context);

auto ft = llvm::FunctionType::get(ti, { ti, ti },
false);

auto f = llvm::Function::Create(ft,
llvm::GlobalValue::ExternalLinkage, "min", m);

f->setCallingConvention(llvm::CallingConv::C);
```

sets C calling convention for the function

Creating a function

```
auto m = llvm::Module::Create(“name”, context);

auto ft = llvm::FunctionType::get(ti, { ti, ti },
false);

auto f = llvm::Function::Create(ft,
llvm::GlobalValue::ExternalLinkage, “min”, m);

f->setCallingConvention(llvm::CallingConv::C);
```

At this point we have created a proper function declaration in LLVM IR. The function can be called in the module, but it does not have any code in it.

Creating code

```
auto args = f->arg_begin();
llvm::Value * first = &* args++;
llvm::Value * second = &* args;

auto b = llvm::BasicBlock::Create(context, “first”, f);
auto cmp = new llvm::ICmpInst(*b,
llvm::ICmpInst::ICMP_SLT, first, second);

auto lt = llvm::BasicBlock::Create(context, “lt”, f);
auto gte = llvm::BasicBlock::Create(context, “gte”, f);

llvm::BranchInst::Create(lt, gte, cmp, b);

llvm::ReturnInst::Create(context, lt, first);

llvm::ReturnInst::Create(context, gte, second);
```

Creating code

Create shorthand values for function arguments

```
auto args = f->arg_begin();
llvm::Value * first = &* args++;
llvm::Value * second = &* args;

auto b = llvm::BasicBlock::Create(context, “first”, f);
auto cmp = new llvm::ICmpInst(*b,
llvm::ICmpInst::ICMP_SLT, first, second);

auto lt = llvm::BasicBlock::Create(context, “lt”, f);
auto gte = llvm::BasicBlock::Create(context, “gte”, f);

llvm::BranchInst::Create(lt, gte, cmp, b);

llvm::ReturnInst::Create(context, lt, first);

llvm::ReturnInst::Create(context, gte, second);
```

Creating code

Create the first basic block

```
o args = f->arg_begin();
m::Value * first = &* args++;
llvm::Value * second = &* args;

auto b = llvm::BasicBlock::Create(context, “first”, f);
auto cmp = new llvm::ICmpInst(*b,
llvm::ICmpInst::ICMP_SLT, first, second);

auto lt = llvm::BasicBl      Create(context, “lt”, f);
auto gte = llvm                                          );

llvm::BranchIns

llvm::ReturnInst::Create(context, lt, first);

llvm::ReturnInst::Create(context, gte, second);
```

Insert signed less than comparison of the arguments at the end of the basic block

Creating code

```
auto args = f->arg_begin();
llvm::Value * first = &* args++;
```

Create basic blocks for i < j and and i >= j cases

```
", f);
```

```
auto lt = llvm::BasicBlock::Create(context, "lt", f);
auto gte = llvm::BasicBlock::Create(context, "gte", f);
llvm::BranchInst::Create(lt, gte, cmp, b);

llvm::ReturnInst

llvm::ReturnInst
```

Conditional branch based on the result of the comparison to either lt, or gte basic blocks

Creating code

```
auto args = f->arg_begin();
llvm::Value * first = &* args++;
llvm::Value

auto b = ll                             ext, “first”, f);
auto cmp =
llvm::ICmpI                             d);

auto lt = llvm::     lock::Create(context, “lt”, f);
auto gte = llvm::     cBlock::Create(context, “gte”, f);

llvm::BranchInst::   eate(lt, gte, cmp, b);

llvm::ReturnInst::Create(context, lt, first);

llvm::ReturnInst::Create(context, gte, second);
```

Return the respective values in the branches

Creating code

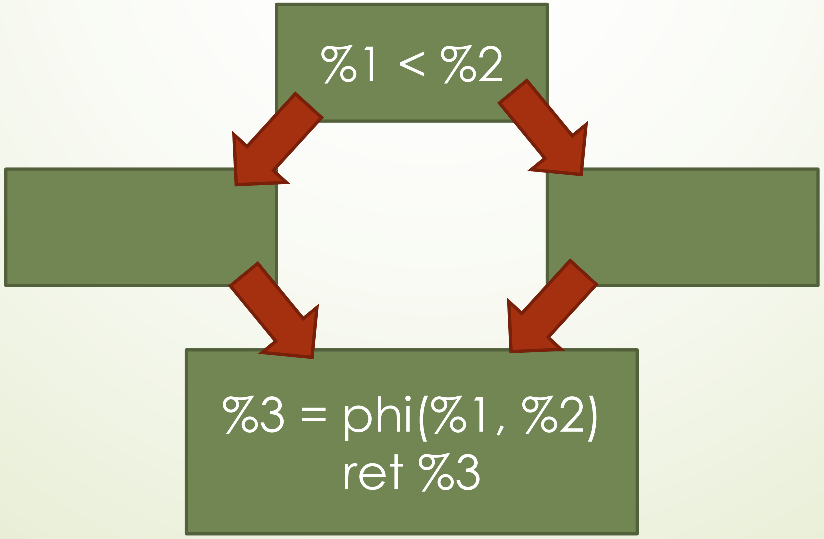
```
define i32 @min(i32 %1, i32 %2) {
    %3 = icmp slt i32 %1, %2
    br i1 %3, label %4, label %5

; <label>:4
    ret i32 %1

; <label>:5
    ret i32 %2
}
```

Working with PHI nodes

- In the min function, we can create only one exit point from the function
- But it would require us to use a PHI node as we would return either first, or second argument in the last basic block based on where we arrived from



Working with PHI nodes

```
define i32 @min(i32 %1, i32 %2) {
    %3 = icmp slt i32 %1, %2
    br i1 %3, label %4, label %5

; <label>:4
    br label %6
; <label>:5
    br label %6
; <label>:6
    %7 = phi i32 [ %1 %4 ], [ %2 %5 ]
    ret i32 %7
}
```

Working with PHI nodes

```
define i32 @min(i32 %1, i32 %2) {
    %3 = icmp slt i32 %1, %2
    br i1 %3, label %4, label %5

; <label>:4
    br label %6
; <label>:5
    br label %6
; <label>:6
    %7 = phi i32 [ %1 %4 ], [ %2 %5 ]
    ret i32 %7
}
```

If arriving from block %4, value will be %1

If arriving from block %5, value will be %2

Working with PHI nodes

```
auto args = f->arg_begin();
llvm::Value * first = args++;
llvm::Value * second = args;

auto b = llvm::BasicBlock::Create(context, “first”, f);
auto cmp = new llvm::ICmpInst(*b,
llvm::ICmpInst::ICMP_SLT, first, second);

auto lt = llvm::BasicBlock::Create(context, “lt”, f);
auto gte = llvm::BasicBlock::Create(context, “gte”, f);
auto end = llvm::BasicBlock::Create(context, “end”, f);

llvm::BranchInst::Create(lt, gte, cmp, b);
llvm::BranchInst::Create(end, lt);
llvm::BranchInst::Create(end, gte);

auto phi = llvm::PHINode::Create(ti, 2, “”, end);
phi->addIncomming(first, lt);
phi->addIncomming(second, gte);

llvm::ReturnInst::Create(context, end, phi);
```

Working with PHI nodes

```
auto args = f->arg_begin();
```

Create the ending basic block and jumps from lt and gte blocks to it (they will be empty)

```
auto gte = llvm::BasicBlock::Create(context, "gte", f);
auto end = llvm::BasicBlock::Create(context, "end", f);

llvm::BranchInst::Create(lt, gte, cmp, b);
llvm::BranchInst::Create(end, lt);
llvm::BranchInst::Create(end, gte);

auto phi = llvm::PHINode::Create(ti, 2, "", end);
phi->addIncomming(first, lt);
phi->addIncomming(second, gte);

llvm::ReturnInst::Create(context, end, phi);
```

Working with PHI nodes

```
auto args = f->arg_begin();
llvm::Value * first = args++;
llvm::Value * second = args;

auto b = llvm::BasicBlock::Create(context, “first”, f);
auto cmp = new llvm::ICmpInst(*b,
llvm::ICmpInst::ICMP_SLT, first, second);

auto lt = llvm::BasicBlock::Create(context, “lt”, f);
auto gte = llvm::BasicBlock::Create(context, “gte”, f);
ate(context, “end”, f);

, cmp, b);
);
e);

auto phi = llvm::PHINode::Create(ti, 2, “”, end);
phi->addIncomming(first, lt);
phi->addIncomming(second, gte);

llvm::ReturnInst::Create(context, end, phi);
```

Create phi node in the last basic block, reserve 2 incoming edges

Add incoming edges

LLVM Toolchain Command Line

Exporting bitcode

```
clang -S -emit-llvm test.c -o test.ll
```

- emits bitcode in human readable form, created by clang frontend from given file
- Slightly more complex than the examples we have seen so far
  - Function attributes
  - Metadata
  - Beware of C++ constructs, name mangling, etc.

Optimizing bitcode

```
opt test.ll -o test.opt.ll -S -mem2reg -dce -constprop
```

- Invokes llvm optimizer on given bitcode file, runs specified passes and outputs the resulting bitcode

```
opt --help
```

- Shows all passes opt understands

C++ API help

```
llc -march=cpp test.ll -o test.ll.c
```

- Takes bitcode in test.ll and compiles it using specified target
- The cpp target is very useful as its output is the C++ API calls required to create the input bitcode
- Definitely more complex than what we have seen so far, but fairly human readable for small examples

Q & A