Code Generation

Dr. Phung Nguyen

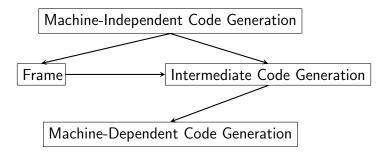
Faculty of Computer Science and Engineering University of Technology HCMC Vietnam National University

November 25, 2020

Outline

Translation to a stack-based machine

Code Generation Design



3/15

Machine-Dependent Code Generation

- Generating specified machine code
 E.g.: emitLDC(20) → "ldc 20"
- Implemented in JasminCode

• Depend on both language and machine

- Depend on both language and machine
- Select instructions

- Depend on both language and machine
- Select instructions

emitREADVAR



remitILOAD.

Depend on both language and machine
 Select instructions
 emitREADVAR

emitFLOAD

Depend on both language and machine
 Select instructions
 emitREADVAR

emitFLOAD

• Select data objects

Depend on both language and machine
 Select instructions
 emitREADVAR(a)
 emitFLOAD(index)

• Select data objects

Depend on both language and machine
 Select instructions
 emitREADVAR(a)
 emitFLOAD(index)

• Select data objects

- Depend on both language and machine
 Select instructions
 emitREADVAR(a)
 emitFLOAD(index)
- Select data objects
- Simulate the execution of the machine

- Depend on both language and machine
 Select instructions
 emitREADVAR(a)
 emitFLOAD(index)
- Select data objects
- Simulate the execution of the machine
 d emitICONST → push()

- Depend on both language and machine
 Select instructions
 emitREADVAR(a)
 emitFLOAD(index)
- Select data objects
- Simulate the execution of the machine
 - \triangleleft emitICONST \rightarrow push()
 - \triangleleft emitISTORE \rightarrow pop()

- Depend on both language and machine ≁emitILOAD(index) Select instructions emitREADVAR(a) \rightarrow emitFLOAD(index)
- Select data objects
- Simulate the execution of the machine
 - \triangleleft emitICONST \rightarrow push()
 - \triangleleft emitISTORE \rightarrow pop()
- Implemented in class Emitter

emitVAR(self,index, varName, inType, fromLabel, toLabel)
 .var 0 is this Lio; from Label0 to Label1

- emitVAR(self,index, varName, inType, fromLabel, toLabel)
 .var 0 is this Lio; from Label0 to Label1
- emitATTRIBUTE(self, lexeme, inType, isFinal, value = None)
 field public static writer Ljava/io/Writer;

- emitVAR(self,index, varName, inType, fromLabel, toLabel)
 .var 0 is this Lio; from Label0 to Label1
- emitATTRIBUTE(self, lexeme, inType, isFinal, value = None)
 field public static writer Ljava/io/Writer;
- emitMETHOD(self, lexeme, inType, isStatic)
 .method public foo(I)I

- emitVAR(self,index, varName, inType, fromLabel, toLabel)
 .var 0 is this Lio; from Label0 to Label1
- emitATTRIBUTE(self, lexeme, inType, isFinal, value = None)
 field public static writer Ljava/io/Writer;
- emitMETHOD(self, lexeme, inType, isStatic)
 .method public foo(I)I
- emitENDMETHOD(self, frame)
 .limit stack 1
 .limit locals 1
 .end method

- emitVAR(self,index, varName, inType, fromLabel, toLabel)
 .var 0 is this Lio; from Label0 to Label1
- emitATTRIBUTE(self, lexeme, inType, isFinal, value = None)
 field public static writer Ljava/io/Writer;
- emitMETHOD(self, lexeme, inType, isStatic)
 .method public foo(I)I
- emitENDMETHOD(self, frame)
 .limit stack 1
 .limit locals 1
 .end method
- emitPROLOG(self, name, parent)
 .source io.java
 .class public io
 .super java/lang/Object

- emitVAR(self,index, varName, inType, fromLabel, toLabel)
 .var 0 is this Lio; from Label0 to Label1
- emitATTRIBUTE(self, lexeme, inType, isFinal, value = None)
 .field public static writer Ljava/io/Writer;
- emitMETHOD(self, lexeme, inType, isStatic)
 .method public foo(I)I
- emitENDMETHOD(self, frame)

 limit stack 1
 limit locals 1
 end method
- emitPROLOG(self, name, parent)
 .source io.java
 .class public io
 .super java/lang/Object
- emitEPILOG(self)

Type

- class IntType(Type)
- class FloatType(Type)
- class StringType(Type)
- class VoidType(Type)
- class BoolType(Type)
- class ClassType(Type): # cname:str
- class ArrayType(Type): # eleType:Type,dimen:List[int]
- class MType(Type): # partype:List[Type],rettype:Type

• emitADDOP(self, lexeme, inType, frame) \Rightarrow iadd, fadd, isub, fsub

- emitADDOP(self, lexeme, inType, frame) \Rightarrow iadd, fadd, isub, fsub
- emitMULOP(self, lexeme, inType, frame) ⇒ imul, fmul, idiv, fdiv

- ullet emitADDOP(self, lexeme, inType, frame) \Rightarrow iadd, fadd, isub, fsub
- emitMULOP(self, lexeme, inType, frame) ⇒ imul, fmul, idiv, fdiv
- emitDIV(self, frame) ⇒ idiv

- emitADDOP(self, lexeme, inType, frame) \Rightarrow iadd, fadd, isub, fsub
- emitMULOP(self, lexeme, inType, frame) ⇒ imul, fmul, idiv, fdiv
- emitDIV(self, frame) ⇒ idiv
- emitMOD(self, frame) ⇒ irem

- emitADDOP(self, lexeme, inType, frame) \Rightarrow iadd, fadd, isub, fsub
- emitMULOP(self, lexeme, inType, frame) ⇒ imul, fmul, idiv, fdiv
- emitDIV(self, frame) ⇒ idiv
- emitMOD(self, frame) ⇒ irem
- emitANDOP(self, frame) \Rightarrow iand

- emitADDOP(self, lexeme, inType, frame) \Rightarrow iadd, fadd, isub, fsub
- emitMULOP(self, lexeme, inType, frame) ⇒ imul, fmul, idiv, fdiv
- emitDIV(self, frame) ⇒ idiv
- emitMOD(self, frame) ⇒ irem
- emitANDOP(self, frame) \Rightarrow iand
- emitOROP(self, frame) ⇒ ior

- emitADDOP(self, lexeme, inType, frame) \Rightarrow iadd, fadd, isub, fsub
- emitMULOP(self, lexeme, inType, frame) ⇒ imul, fmul, idiv, fdiv
- emitDIV(self, frame) ⇒ idiv
- emitMOD(self, frame) ⇒ irem
- emitANDOP(self, frame) \Rightarrow iand
- emitOROP(self, frame) \Rightarrow ior
- emitREOP(self, op, inType, frame) \Rightarrow code for >, < >=, <=, !=, ==

- emitADDOP(self, lexeme, inType, frame) \Rightarrow iadd, fadd, isub, fsub
- emitMULOP(self, lexeme, inType, frame) ⇒ imul, fmul, idiv, fdiv
- emitDIV(self, frame) ⇒ idiv
- emitMOD(self, frame) ⇒ irem
- emitANDOP(self, frame) \Rightarrow iand
- emitOROP(self, frame) ⇒ ior
- emitREOP(self, op, inType, frame) \Rightarrow code for >, < >=, <=, !=, ==
- ullet emitRELOP(self, op, inType, trueLabel, falseLabel, frame) \Rightarrow code for condition in if statement

• emitREADVAR(self, name, inType, index, frame) \Rightarrow [ifa]load

- emitREADVAR(self, name, inType, index, frame) ⇒ [ifa]load
- emitALOAD(self, inType, frame)) \Rightarrow [ifa]aload

- emitREADVAR(self, name, inType, index, frame) ⇒ [ifa]load
- emitALOAD(self, inType, frame)) ⇒ [ifa]aload
- emitWRITEVAR(self, name, inType, index, frame) \Rightarrow [ifa]store

- emitREADVAR(self, name, inType, index, frame) ⇒ [ifa]load
- emitALOAD(self, inType, frame)) ⇒ [ifa]aload
- emitWRITEVAR(self, name, inType, index, frame) \Rightarrow [ifa]store
- emitASTORE(self, inType, frame) \Rightarrow [ifa]astore

- emitREADVAR(self, name, inType, index, frame) ⇒ [ifa]load
- emitALOAD(self, inType, frame)) ⇒ [ifa]aload
- emitWRITEVAR(self, name, inType, index, frame) \Rightarrow [ifa]store
- emitASTORE(self, inType, frame) \Rightarrow [ifa]astore
- emitGETSTATIC(self, lexeme, inType, frame) ⇒ getstatic

- emitREADVAR(self, name, inType, index, frame) ⇒ [ifa]load
- emitALOAD(self, inType, frame)) ⇒ [ifa]aload
- emitWRITEVAR(self, name, inType, index, frame) \Rightarrow [ifa]store
- emitASTORE(self, inType, frame) \Rightarrow [ifa]astore
- emitGETSTATIC(self, lexeme, inType, frame) ⇒ getstatic
- ullet emitGETFIELD(self, lexeme, inType, frame) \Rightarrow getfield

Read/Write Variables APIs

- emitREADVAR(self, name, inType, index, frame) ⇒ [ifa]load
- $\bullet \ \mathsf{emitALOAD}(\mathsf{self}, \ \mathsf{inType}, \ \mathsf{frame})) \Rightarrow [\mathsf{ifa}] \mathsf{aload}$
- $\bullet \ \, \mathsf{emitWRITEVAR}(\mathsf{self}, \ \mathsf{name}, \ \mathsf{inType}, \ \mathsf{index}, \ \mathsf{frame}) \Rightarrow [\mathsf{ifa}] \mathsf{store}$
- emitASTORE(self, inType, frame) \Rightarrow [ifa]astore
- emitGETSTATIC(self, lexeme, inType, frame) ⇒ getstatic
- ullet emitGETFIELD(self, lexeme, inType, frame) \Rightarrow getfield
- $\bullet \ \, \mathsf{emitPUTSTATIC}(\mathsf{self}, \, \mathsf{lexeme}, \, \mathsf{inType}, \, \mathsf{frame}) \Rightarrow \mathsf{putstatic}$

Read/Write Variables APIs

- emitREADVAR(self, name, inType, index, frame) ⇒ [ifa]load
- emitALOAD(self, inType, frame)) ⇒ [ifa]aload
- emitWRITEVAR(self, name, inType, index, frame) \Rightarrow [ifa]store
- emitASTORE(self, inType, frame) \Rightarrow [ifa]astore
- emitGETSTATIC(self, lexeme, inType, frame) ⇒ getstatic
- ullet emitGETFIELD(self, lexeme, inType, frame) \Rightarrow getfield
- $\bullet \ \, \mathsf{emitPUTSTATIC}(\mathsf{self}, \, \mathsf{lexeme}, \, \mathsf{inType}, \, \mathsf{frame}) \Rightarrow \mathsf{putstatic}$
- $\bullet \ \ \mathsf{emitPUTFIELD}(\mathsf{self}, \ \mathsf{lexeme}, \ \mathsf{inType}, \ \mathsf{frame}) \Rightarrow \mathsf{putfield}$

• emitPUSHICONST(self, input, frame) ⇒ iconst, bipush, sipush, ldc

- ullet emitPUSHICONST(self, input, frame) \Rightarrow iconst, bipush, sipush, ldc
- emitPUSHFCONST(self, input, frame) ⇒ fconst, ldc

- ullet emitPUSHICONST(self, input, frame) \Rightarrow iconst, bipush, sipush, ldc
- emitPUSHFCONST(self, input, frame) ⇒ fconst, ldc
- emitINVOKESTATIC(self, lexeme, inType, frame) ⇒ invokestatic

- emitPUSHICONST(self, input, frame) ⇒ iconst, bipush, sipush, ldc
- emitPUSHFCONST(self, input, frame) ⇒ fconst, ldc
- emitINVOKESTATIC(self, lexeme, inType, frame) ⇒ invokestatic
- emitINVOKESPECIAL(self, frame, lexeme=None, inType=None)
 invokespecial

- emitPUSHICONST(self, input, frame) ⇒ iconst, bipush, sipush, ldc
- emitPUSHFCONST(self, input, frame) ⇒ fconst, ldc
- emitINVOKESTATIC(self, lexeme, inType, frame) ⇒ invokestatic
- emitINVOKESPECIAL(self, frame, lexeme=None, inType=None)
 invokespecial
- emitINVOKEVIRTUAL(self, lexeme, inType, frame) ⇒ invokevirtual

- emitPUSHICONST(self, input, frame) ⇒ iconst, bipush, sipush, ldc
- emitPUSHFCONST(self, input, frame) ⇒ fconst, ldc
- emitINVOKESTATIC(self, lexeme, inType, frame) ⇒ invokestatic
- emitINVOKESPECIAL(self, frame, lexeme=None, inType=None)
 invokespecial
- emitINVOKEVIRTUAL(self, lexeme, inType, frame) ⇒ invokevirtual
- emitIFTRUE(self, label, frame) ⇒ ifgt

- emitPUSHICONST(self, input, frame) ⇒ iconst, bipush, sipush, ldc
- emitPUSHFCONST(self, input, frame) ⇒ fconst, ldc
- emitINVOKESTATIC(self, lexeme, inType, frame) ⇒ invokestatic
- emitINVOKESPECIAL(self, frame, lexeme=None, inType=None)
 invokespecial
- emitINVOKEVIRTUAL(self, lexeme, inType, frame) ⇒ invokevirtual
- emitIFTRUE(self, label, frame) ⇒ ifgt
- ullet emitIFFALSE(self, label, frame) \Rightarrow ifle

- ullet emitPUSHICONST(self, input, frame) \Rightarrow iconst, bipush, sipush, ldc
- emitPUSHFCONST(self, input, frame) ⇒ fconst, ldc
- emitINVOKESTATIC(self, lexeme, inType, frame) ⇒ invokestatic
- emitINVOKESPECIAL(self, frame, lexeme=None, inType=None)
 invokespecial
- emitINVOKEVIRTUAL(self, lexeme, inType, frame) ⇒ invokevirtual
- emitIFTRUE(self, label, frame) ⇒ ifgt
- ullet emitIFFALSE(self, label, frame) \Rightarrow ifle
- emitDUP(self,frame) ⇒ dup

- ullet emitPUSHICONST(self, input, frame) \Rightarrow iconst, bipush, sipush, ldc
- emitPUSHFCONST(self, input, frame) ⇒ fconst, ldc
- emitINVOKESTATIC(self, lexeme, inType, frame) ⇒ invokestatic
- emitINVOKESPECIAL(self, frame, lexeme=None, inType=None)
 invokespecial
- emitINVOKEVIRTUAL(self, lexeme, inType, frame) ⇒ invokevirtual
- emitIFTRUE(self, label, frame) ⇒ ifgt
- ullet emitIFFALSE(self, label, frame) \Rightarrow ifle
- emitDUP(self,frame) ⇒ dup
- emitPOP(self,frame) ⇒ pop

- ullet emitPUSHICONST(self, input, frame) \Rightarrow iconst, bipush, sipush, ldc
- emitPUSHFCONST(self, input, frame) ⇒ fconst, ldc
- emitINVOKESTATIC(self, lexeme, inType, frame) ⇒ invokestatic
- emitINVOKESPECIAL(self, frame, lexeme=None, inType=None)
 invokespecial
- emitINVOKEVIRTUAL(self, lexeme, inType, frame) ⇒ invokevirtual
- emitIFTRUE(self, label, frame) ⇒ ifgt
- ullet emitIFFALSE(self, label, frame) \Rightarrow ifle
- emitDUP(self,frame) ⇒ dup
- emitPOP(self,frame) ⇒ pop
- emitl2F(self, frame) ⇒ i2f

- ullet emitPUSHICONST(self, input, frame) \Rightarrow iconst, bipush, sipush, ldc
- emitPUSHFCONST(self, input, frame) \Rightarrow fconst, ldc
- emitINVOKESTATIC(self, lexeme, inType, frame) ⇒ invokestatic
- emitINVOKESPECIAL(self, frame, lexeme=None, inType=None)
 invokespecial
- emitINVOKEVIRTUAL(self, lexeme, inType, frame) ⇒ invokevirtual
- emitIFTRUE(self, label, frame) ⇒ ifgt
- emitIFFALSE(self, label, frame) \Rightarrow ifle
- emitDUP(self,frame) ⇒ dup
- emitPOP(self,frame) ⇒ pop
- emitl2F(self, frame) \Rightarrow i2f
- emitRETURN(self, inType, frame) ⇒ return, ireturn

- ullet emitPUSHICONST(self, input, frame) \Rightarrow iconst, bipush, sipush, ldc
- emitPUSHFCONST(self, input, frame) ⇒ fconst, ldc
- emitINVOKESTATIC(self, lexeme, inType, frame) ⇒ invokestatic
- emitINVOKESPECIAL(self, frame, lexeme=None, inType=None)
 invokespecial
- emitINVOKEVIRTUAL(self, lexeme, inType, frame) ⇒ invokevirtual
- emitIFTRUE(self, label, frame) ⇒ ifgt
- ullet emitIFFALSE(self, label, frame) \Rightarrow ifle
- emitDUP(self,frame) ⇒ dup
- emitPOP(self,frame) ⇒ pop
- emitl2F(self, frame) ⇒ i2f
- emitRETURN(self, inType, frame) ⇒ return, ireturn
- emitLABEL(self, label, frame) \Rightarrow Label



- ullet emitPUSHICONST(self, input, frame) \Rightarrow iconst, bipush, sipush, ldc
- emitPUSHFCONST(self, input, frame) ⇒ fconst, ldc
- emitINVOKESTATIC(self, lexeme, inType, frame) ⇒ invokestatic
- emitINVOKESPECIAL(self, frame, lexeme=None, inType=None)
 invokespecial
- emitINVOKEVIRTUAL(self, lexeme, inType, frame) ⇒ invokevirtual
- emitIFTRUE(self, label, frame) ⇒ ifgt
- ullet emitIFFALSE(self, label, frame) \Rightarrow ifle
- emitDUP(self,frame) ⇒ dup
- emitPOP(self,frame) ⇒ pop
- emitl2F(self, frame) ⇒ i2f
- emitRETURN(self, inType, frame) ⇒ return, ireturn
- emitLABEL(self, label, frame) \Rightarrow Label
- emitGOTO(self, label, frame) ⇒ goto

- Labels: are valid in the body of a method
 - ⊲ getNewLabel(): return a new label

- Labels: are valid in the body of a method
 - ⊲ getNewLabel(): return a new label
 - ⊲ getStartLabel(): return the beginning label of a scope

- Labels: are valid in the body of a method
 - ⊲ getNewLabel(): return a new label
 - ⊲ getStartLabel(): return the beginning label of a scope
 - ⊲ getEndLabel(): return the end label of a scope

- Labels: are valid in the body of a method
 - ⊲ getNewLabel(): return a new label
 - ⊲ getStartLabel(): return the beginning label of a scope
 - ⊲ getEndLabel(): return the end label of a scope
 - ⊲ getContinueLabel(): return the label where a continue should come

- Labels: are valid in the body of a method
 - ⊲ getNewLabel(): return a new label
 - ⊲ getStartLabel(): return the beginning label of a scope
 - ⊲ getEndLabel(): return the end label of a scope
 - ⊲ getContinueLabel(): return the label where a continue should come
 - \triangleleft getBreakLabel(): return the label where a break should come

- Labels: are valid in the body of a method
 - ⊲ getNewLabel(): return a new label
 - ⊲ getStartLabel(): return the beginning label of a scope
 - \triangleleft getEndLabel(): return the end label of a scope
 - ⊲ getContinueLabel(): return the label where a continue should come
 - \triangleleft getBreakLabel(): return the label where a break should come
 - d enterScope()

- Labels: are valid in the body of a method
 - ⊲ getNewLabel(): return a new label
 - ⊲ getStartLabel(): return the beginning label of a scope
 - ⊲ getEndLabel(): return the end label of a scope
 - ⊲ getContinueLabel(): return the label where a continue should come
 - ⊲ getBreakLabel(): return the label where a break should come
 - ⊲ enterScope()
 - ⊲ exitScope()

- Labels: are valid in the body of a method
 - ⊲ getNewLabel(): return a new label
 - ⊲ getStartLabel(): return the beginning label of a scope
 - ⊲ getEndLabel(): return the end label of a scope
 - ⊲ getContinueLabel(): return the label where a continue should come
 - ⊲ getBreakLabel(): return the label where a break should come
 - ⊲ enterScope()
 - < exitScope()</pre>
 - ⊲ enterLoop()

- Labels: are valid in the body of a method
 - ⊲ getNewLabel(): return a new label
 - ⊲ getStartLabel(): return the beginning label of a scope
 - ⊲ getEndLabel(): return the end label of a scope
 - ⊲ getContinueLabel(): return the label where a continue should come
 - ⊲ getBreakLabel(): return the label where a break should come
 - ⊲ enterScope()
 - < exitScope()</pre>
 - ⊲ enterLoop()
 - ⊲ exitLoop()

Local variable array

- Local variable array
 - ⊲ getNewIndex(): return a new index for a variable

- Local variable array
 - □ getNewIndex(): return a new index for a variable
 - ⊲ getMaxIndex(): return the size of the local variable array

- Local variable array
 - ⊲ getNewIndex(): return a new index for a variable
 - ⊲ getMaxIndex(): return the size of the local variable array
- Operand stack

- Local variable array
 - ⊲ getNewIndex(): return a new index for a variable
 - ⊲ getMaxIndex(): return the size of the local variable array
- Operand stack
 - q push(): simulating a push execution

- Local variable array
 - ⊲ getNewIndex(): return a new index for a variable
 - ⊲ getMaxIndex(): return the size of the local variable array
- Operand stack
 - q push(): simulating a push execution
 - ⊲ pop(): simulating a pop execution

- Local variable array
 - ⊲ getNewIndex(): return a new index for a variable
 - ⊲ getMaxIndex(): return the size of the local variable array
- Operand stack
 - push(): simulating a push execution
 - op(): simulating a pop execution
 - \triangleleft getMaxOpStackSize(): return the max size of the operand stack

- Local variable array
 - ⊲ getNewIndex(): return a new index for a variable
 - ⊲ getMaxIndex(): return the size of the local variable array
- Operand stack
 - push(): simulating a push execution
 - op(): simulating a pop execution
 - ⊲ getMaxOpStackSize(): return the max size of the operand stack
- Implemented in class Frame

Machine-Independent Code Generation

- Based on the source language
- Use facilities of Frame and Intermediate Code Generation (Emitter)

BKIT-Java mapping

- A source program \Rightarrow Java class
- A global variable ⇒ a static field
- A function \Rightarrow a static method
- ullet A parameter \Rightarrow a parameter
- A local variable \Rightarrow a local variable
- An expression \Rightarrow an expression
- ullet A statement \Rightarrow a statement
- An invocation \Rightarrow an invocation

• Use BCEL to know which code should be generated

- Use BCEL to know which code should be generated
- Generate code for expressions first

- Use BCEL to know which code should be generated
- Generate code for expressions first
- Generate code for statements later

- Use BCEL to know which code should be generated
- Generate code for expressions first
- Generate code for statements later
- Good luck