# CS738: Advanced Compiler Optimizations Types and Program Analysis

#### Amey Karkare

karkare@cse.iitk.ac.in

http://www.cse.iitk.ac.in/~karkare/cs738

Department of CSE, IIT Kanpur



#### Reference Book

Types and Programming Languages by Benjamin C. Pierce

#### Type: Definition

#### type /t∧ɪp/ •

#### noun

- a category of people or things having common characteristics.
   "this type of heather grows better in a drier habitat"
   synonyms: kind, sort, variety, class, category, classification, group, set, bracket, genre, genus,
   species, family, order, breed, race, strain; More
- 2. a person or thing exemplifying the ideal or defining characteristics of something. "she characterized his witty sayings as the type of modern wisdom" synonyms: epitome, quintessence, essence, perfect example, archetype, model, pattern, paradigm, exemplar, embodiment, personification, avatar; prototype "she characterized his witty sayings as the type of modern wisdom"

### Types in Programming

► A collection of *values* 



#### Types in Programming

► A collection of *values* 



► The operations that are permitted on these values

#### Type System

A collection of rules for checking the correctness of usages of types

#### Type System

- A collection of rules for checking the correctness of usages of types
  - "Consistency" of programs

Typed

- Typed
  - ► C, C++, Java, Python, ...

- Typed
  - ► C, C++, Java, Python, . . .
- Untyped

- Typed
  - ► C, C++, Java, Python, . . .
- Untyped
  - Assembly, any other?

	Statically Typed	Dynamically Typed
Strongly Typed		
Weekly Typed		

	Statically Typed	Dynamically Typed
Strongly Typed	ML, Haskell, Pascal (almost), Java (almost)	
Weekly Typed		

	Statically Typed	Dynamically Typed
Strongly Typed	ML, Haskell, Pascal (almost), Java (almost)	Lisp, Scheme
Weekly Typed		

	Statically Typed	Dynamically Typed
Strongly Typed	ML, Haskell, Pascal (almost), Java (almost)	
Weekly Typed	C, C++	

	Statically Typed	Dynamically Typed
Strongly Typed	ML, Haskell, Pascal (almost), Java (almost)	Lisp, Scheme
Weekly Typed	C, C++	Perl

Error Detection

- Error Detection
  - Language Safety

- Error Detection
  - Language Safety
  - Verification

- Error Detection
  - Language Safety
  - Verification
- Abstraction

- Error Detection
  - Language Safety
  - Verification
- Abstraction
- Documentation

- Error Detection
  - Language Safety
  - Verification
- Abstraction
- Documentation
- Maintenance

- Error Detection
  - Language Safety
  - Verification
- Abstraction
- Documentation
- Maintenance
- Efficiency

$$t := -terms$$

$$t := -terms - constant true$$

```
\begin{array}{ccc} t := & & -\textit{terms} \\ & \text{true} & -\textit{constant true} \\ & \text{false} & -\textit{constant false} \end{array}
```

```
\begin{array}{ccc} t := & & -\textit{terms} \\ & \text{true} & -\textit{constant true} \\ & \text{false} & -\textit{constant false} \\ & \text{if } t \text{ then } t \text{ else } t & -\textit{conditional} \end{array}
```

```
\begin{array}{cccc} t := & & -\textit{terms} \\ & \text{true} & -\textit{constant true} \\ & \text{false} & -\textit{constant false} \\ & \text{if } t \text{ then } t \text{ else } t & -\textit{conditional} \\ & 0 & -\textit{constant zero} \end{array}
```

```
\begin{array}{lll} t \coloneqq & -\textit{terms} \\ & \texttt{true} & -\textit{constant true} \\ & \texttt{false} & -\textit{constant false} \\ & \texttt{if } t \texttt{ then } t \texttt{ else } t & -\textit{conditional} \\ & 0 & -\textit{constant zero} \\ & \texttt{succ } t & -\textit{successor} \end{array}
```

```
t:= - terms
    true - constant true
    false - constant false
    if t then t else t - conditional
    o - constant zero
    succ t - successor
    pred t - predecessor
```

```
t :=
                                 terms

    constant true

     true

constant false

      false
      if t then t else t
                                 conditional
                                 constant zero
     O
      succ t
                                 - successor
                                 - predecessor
     pred t
      iszero t

    zero test
```

The set of *terms* is the smallest set  $\mathcal{T}$  such that

The set of *terms* is the smallest set T such that

1.  $\{\text{true}, \text{false}, 0\} \subseteq \mathcal{T}$ 

#### The set of *terms* is the smallest set $\mathcal{T}$ such that

- 1.  $\{true, false, 0\} \subseteq \mathcal{T}$
- 2. if  $t_1 \in \mathcal{T},$  then  $\{\text{succ } t_1, \text{pred } t_1, \text{iszero } t_1\} \subseteq \mathcal{T}$

#### The set of *terms* is the smallest set $\mathcal{T}$ such that

- 1.  $\{true, false, 0\} \subseteq \mathcal{T}$
- 2. if  $t_1 \in \mathcal{T}$ , then  $\{\text{succ } t_1, \text{pred } t_1, \text{iszero } t_1\} \subseteq \mathcal{T}$
- 3. if  $t_1 \in \mathcal{T}, t_2 \in \mathcal{T}$ , and  $t_3 \in \mathcal{T}$  then if  $t_1$  then  $t_2$  else  $t_3 \in \mathcal{T}$

#### Syntax: Inference Rules

The set of  $\textit{terms}, \, \mathcal{T} \,$  is defined by the following rules:

## Syntax: Inference Rules

The set of *terms*,  $\mathcal{T}$  is defined by the following rules:

true  $\in \mathcal{T}$  false  $\in \mathcal{T}$ 

 $0\in \mathcal{T}$ 

## Syntax: Inference Rules

The set of *terms*,  $\mathcal{T}$  is defined by the following rules:

$$\texttt{true} \in \mathcal{T}$$

$$\texttt{false} \in \mathcal{T}$$

$$0\in \mathcal{T}$$

$$\frac{t_1 \in \mathcal{T}}{\text{succ } t_1 \in \mathcal{T}}$$

$$\frac{t_1 \in \mathcal{T}}{\text{pred } t_1 \in \mathcal{T}}$$

$$\frac{t_1 \in \mathcal{T}}{\text{pred } t_1 \in \mathcal{T}} \qquad \frac{t_1 \in \mathcal{T}}{\text{iszero } t_1 \in \mathcal{T}}$$

## Syntax: Inference Rules

The set of *terms*,  $\mathcal{T}$  is defined by the following rules:

$$\begin{array}{ll} \text{true} \in \mathcal{T} & \text{false} \in \mathcal{T} & \textbf{0} \in \mathcal{T} \\ \\ \frac{t_1 \in \mathcal{T}}{\text{succ } t_1 \in \mathcal{T}} & \frac{t_1 \in \mathcal{T}}{\text{pred } t_1 \in \mathcal{T}} & \frac{t_1 \in \mathcal{T}}{\text{iszero } t_1 \in \mathcal{T}} \\ \\ & \frac{t_1 \in \mathcal{T} \quad t_2 \in \mathcal{T} \quad t_3 \in \mathcal{T}}{\text{if } t_1 \text{ then } t_2 \text{ else } t_3 \in \mathcal{T}} \end{array}$$

$$S_0 = \emptyset$$

$$\mathcal{S}_0 = \emptyset$$
  $\mathcal{S}_{i+1} = \{\text{true}, \text{false}, 0\}$ 

```
 \begin{array}{rcl} \mathcal{S}_0 & = & \emptyset \\ \\ \mathcal{S}_{i+1} & = & \left\{ \texttt{true}, \texttt{false}, 0 \right\} \\ & & \cup \left\{ \texttt{succ} \ t_1, \texttt{pred} \ t_1, \texttt{iszero} \ t_1 \ \middle| \ t_1 \in \mathcal{S}_i \right\} \\ \end{array}
```

```
 \begin{split} \mathcal{S}_0 &= \emptyset \\ \mathcal{S}_{i+1} &= \{\texttt{true}, \texttt{false}, 0\} \\ & \cup \{\texttt{succ}\, t_1, \texttt{pred}\, t_1, \texttt{iszero}\, t_1 \mid t_1 \in \mathcal{S}_i\} \\ & \cup \{\texttt{if}\, t_1 \, \texttt{then}\, t_2 \, \texttt{else}\, t_3 \mid t_1, t_2, t_2 \in \mathcal{S}_i\} \end{split}
```

$$\begin{array}{rcl} \mathcal{S}_0 &=& \emptyset \\ \mathcal{S}_{i+1} &=& \{\texttt{true}, \texttt{false}, 0\} \\ && \cup \{\texttt{succ} \ t_1, \texttt{pred} \ t_1, \texttt{iszero} \ t_1 \mid t_1 \in \mathcal{S}_i\} \\ && \cup \{\texttt{if} \ t_1 \ \texttt{then} \ t_2 \ \texttt{else} \ t_3 \mid t_1, t_2, t_2 \in \mathcal{S}_i\} \end{array}$$
 Let  $\mathcal{S} = \bigcup_i \mathcal{S}_i$ .

$$\begin{array}{rcl} \mathcal{S}_0 &=& \emptyset \\ \mathcal{S}_{i+1} &=& \{\texttt{true}, \texttt{false}, 0\} \\ && \cup \{\texttt{succ}\ t_1, \texttt{pred}\ t_1, \texttt{iszero}\ t_1 \mid t_1 \in \mathcal{S}_i\} \\ && \cup \{\texttt{if}\ t_1\ \texttt{then}\ t_2\ \texttt{else}\ t_3 \mid t_1, t_2, t_2 \in \mathcal{S}_i\} \end{array}$$
 Let  $\mathcal{S} = \bigcup_i \mathcal{S}_i$ . Then,  $\mathcal{T} = \mathcal{S}$ .

 $\blacktriangleright \ \, \text{Any} \,\, t \in \mathcal{T}$ 

- $\blacktriangleright \ \, \text{Any} \,\, t \in \mathcal{T}$ 
  - ▶ Either a ground term, i.e.  $\in \{ \text{true}, \text{false}, 0 \}$

- ▶ Any  $t \in \mathcal{T}$ 
  - ▶ Either a ground term, i.e. ∈ {true, false, 0}
  - lacktriangle Or is created from some smaller terms  $\in \mathcal{T}$

- ▶ Any  $t \in \mathcal{T}$ 
  - ▶ Either a ground term, i.e. ∈ {true, false, 0}
  - ightharpoonup Or is created from some smaller terms  $\in \mathcal{T}$
- Allows for inductive definitions and inductive proofs.

- ▶ Any  $t \in \mathcal{T}$ 
  - ▶ Either a ground term, i.e. ∈ {true, false, 0}
  - ightharpoonup Or is created from some smaller terms  $\in \mathcal{T}$
- Allows for inductive definitions and inductive proofs.
- Three sample inductive properties

- ▶ Any  $t \in \mathcal{T}$ 
  - ▶ Either a ground term, i.e. ∈ {true, false, 0}
  - ightharpoonup Or is created from some smaller terms  $\in \mathcal{T}$
- Allows for inductive definitions and inductive proofs.
- Three sample inductive properties
  - Consts(t)

- ▶ Any  $t \in \mathcal{T}$ 
  - ▶ Either a ground term, i.e. ∈ {true, false, 0}
  - ightharpoonup Or is created from some smaller terms  $\in \mathcal{T}$
- Allows for inductive definitions and inductive proofs.
- Three sample inductive properties
  - Consts(t)
  - ► size(t)

- ▶ Any  $t \in \mathcal{T}$ 
  - ▶ Either a ground term, i.e. ∈ {true, false, 0}
  - ightharpoonup Or is created from some smaller terms  $\in \mathcal{T}$
- Allows for inductive definitions and inductive proofs.
- Three sample inductive properties
  - Consts(t)
  - size(t)
  - depth(t)

$$Consts(true) = \{true\}$$

```
Consts(true) = \{true\}

Consts(false) = \{false\}
```

```
Consts(true) = \{true\}

Consts(false) = \{false\}

Consts(0) = \{0\}
```

```
Consts(true) = \{true\}
Consts(false) = \{false\}
Consts(0) = \{0\}
Consts(succ t) = Consts(t)
```

```
Consts(true) = \{true\}
Consts(false) = \{false\}
Consts(0) = \{0\}
Consts(succ t) = Consts(t)
Consts(pred t) = Consts(t)
```

```
Consts(true) = \{true\}
Consts(false) = \{false\}
Consts(0) = \{0\}
Consts(succ t) = Consts(t)
Consts(pred t) = Consts(t)
Consts(iszero t) = Consts(t)
```

```
Consts(true) = {true}
              Consts(false) = \{false\}
                    Consts(0) = \{0\}
              Consts(succ t) = Consts(t)
              Consts(pred t) = Consts(t)
            Consts(iszerot) = Consts(t)
Consts(if t_1 then t_2 else t_3) = Consts(t_1)
                                    \cup Consts(t<sub>2</sub>)
                                    \cup Consts(t<sub>3</sub>)
```

$$size(true) = 1$$

$$size(true) = 1$$
  
 $size(false) = 1$ 

$$size(true) = 1$$
  
 $size(false) = 1$   
 $size(0) = 1$ 

```
size(true) = 1

size(false) = 1

size(0) = 1

size(succ t) = size(t) + 1
```

```
size(true) = 1

size(false) = 1

size(0) = 1

size(succt) = size(t) + 1

size(predt) = size(t) + 1
```

```
size(true) = 1

size(false) = 1

size(0) = 1

size(succt) = size(t) + 1

size(predt) = size(t) + 1

size(iszerot) = size(t) + 1
```

```
size(true) = 1

size(false) = 1

size(0) = 1

size(succt) = size(t) + 1

size(predt) = size(t) + 1

size(iszerot) = size(t) + 1

size(if t_1 then t_2 else t_3) = size(t_1) + size(t_2) + size(t_3)
```

### depth

- ▶ The maximum depth of the abstract syntax tree of a term t.
- ▶ Equivalently, the smallest i such that  $t \in S_i$ .

### depth

- ▶ The maximum depth of the abstract syntax tree of a term t.
- ▶ Equivalently, the smallest *i* such that  $t \in S_i$ .

$$depth(true) = 1$$

### depth

- The maximum depth of the abstract syntax tree of a term t.
- ▶ Equivalently, the smallest *i* such that  $t \in S_i$ .

```
depth(true) = 1
depth(false) = 1
```

- The maximum depth of the abstract syntax tree of a term t.
- ▶ Equivalently, the smallest *i* such that  $t \in S_i$ .

```
depth(true) = 1

depth(false) = 1

depth(0) = 1
```

- The maximum depth of the abstract syntax tree of a term t.
- ▶ Equivalently, the smallest *i* such that  $t \in S_i$ .

```
depth(true) = 1

depth(false) = 1

depth(0) = 1

depth(succ t) = depth(t) + 1
```

- The maximum depth of the abstract syntax tree of a term t.
- ▶ Equivalently, the smallest *i* such that  $t \in S_i$ .

```
depth(true) = 1

depth(false) = 1

depth(0) = 1

depth(succ t) = depth(t) + 1

depth(pred t) = depth(t) + 1
```

- The maximum depth of the abstract syntax tree of a term t.
- ▶ Equivalently, the smallest *i* such that  $t \in S_i$ .

```
depth(true) = 1

depth(false) = 1

depth(0) = 1

depth(succt) = depth(t) + 1

depth(predt) = depth(t) + 1

depth(iszerot) = depth(t) + 1
```

- The maximum depth of the abstract syntax tree of a term t.
- ▶ Equivalently, the smallest *i* such that  $t \in S_i$ .

```
\begin{array}{rcl} \textit{depth}(\texttt{true}) &=& 1 \\ \textit{depth}(\texttt{false}) &=& 1 \\ \textit{depth}(0) &=& 1 \\ \textit{depth}(\texttt{succ}\,t) &=& \textit{depth}(t) + 1 \\ \textit{depth}(\texttt{pred}\,t) &=& \textit{depth}(t) + 1 \\ \textit{depth}(\texttt{iszero}\,t) &=& \textit{depth}(t) + 1 \\ \textit{depth}(\texttt{if}\,t_1\,\texttt{then}\,t_2\,\texttt{else}\,t_3) &=& \max(\textit{depth}(t_1),\textit{depth}(t_2),\\ \textit{depth}(t_3)) + 1 \end{array}
```

## A Simple Property of Terms

► The number of distinct constants in a term t is no greater than the size of t.

$$|\textit{Consts}(t)| \leq \textit{size}(t)$$

## A Simple Property of Terms

The number of distinct constants in a term t is no greater than the size of t.

$$|\textit{Consts}(t)| \leq \textit{size}(t)$$

Proof: Exercise.

V :=

values

- values
- value true

```
V := - values \\ true - value true \\ false - value false
```

```
V := - values \ true - value true \ false - value talse \ 0 - value zero
```

ightharpoonup t o t' denotes "t evaluates to t' in one step"

if true then  $t_2$  else  $t_3 \rightarrow t_2$ 

ightharpoonup t o t' denotes "t evaluates to t' in one step"

if true then  $t_2$  else  $t_3 \rightarrow t_2$ 

if false then  $t_2$  else  $t_3 \rightarrow t_3$ 

ightharpoonup t ightharpoonupt denotes "t evaluates to t' in one step"

if true then  $t_2$  else  $t_3 \rightarrow t_2$ 

if false then  $t_2$  else  $t_3 \rightarrow t_3$ 

$$t_1 \rightarrow t_1^\prime$$

if  $t_1$  then  $t_2$  else  $t_3 \to \text{if } t_1'$  then  $t_2$  else  $t_3$ 

$$\frac{t_1 \to t_1'}{\text{succ } t_1 \ \to \text{succ } t_1'}$$

$$\frac{t_1 \to t_1'}{\text{succ } t_1 \ \to \text{succ } t_1'}$$
 
$$\text{pred } 0 \to 0$$

$$\begin{aligned} \frac{t_1 \to t_1'}{\text{succ } t_1 \ \to \text{succ } t_1'} \\ & \text{pred } 0 \to 0 \end{aligned}$$
 
$$\text{pred (succ } v) \to v$$

$$\begin{aligned} \frac{t_1 \rightarrow t_1'}{\text{succ } t_1 \ \rightarrow \text{succ } t_1'} \\ \text{pred } 0 \rightarrow 0 \\ \\ \text{pred } (\text{succ } v) \rightarrow v \\ \\ \frac{t_1 \rightarrow t_1'}{\text{pred } t_1 \ \rightarrow \text{pred } t_1'} \end{aligned}$$

ightharpoonup t o t' denotes "t evaluates to t' in one step"

iszero $0 \rightarrow \text{true}$ 

ightharpoonup t o t' denotes "t evaluates to t' in one step"

iszero $0 \rightarrow \text{true}$ 

 $\texttt{iszero} \, (\texttt{succ} \, v) \to \texttt{false}$ 

ightharpoonup t ightharpoonupt denotes "t evaluates to t' in one step"

iszero 
$$0 \rightarrow \text{true}$$
 iszero (succ  $v$ )  $\rightarrow$  false 
$$\frac{t_1 \rightarrow t_1'}{\text{iszero } t_1 \rightarrow \text{iszero } t_1'}$$

#### **Normal Form**

A term is t in normal form if no evaluation rule applies to it.

#### **Normal Form**

- ▶ A term is t in normal form if no evaluation rule applies to it.
- ▶ In other words, there is no t' such that  $t \to t'$ .

## **Evaluation Sequence**

➤ An evaluation sequence starting from a term t is a (finite or infinite) sequence of terms t<sub>1</sub>, t<sub>2</sub>,..., such that

$$t \rightarrow t_1 \\$$

$$t_1 \rightarrow t_2 \,$$

etc.

#### Stuck Term

A term is said to be **stuck** if it is a normal form but not a value.

#### Stuck Term

- A term is said to be **stuck** if it is a normal form but not a value.
- ► A simple notion of "run-time type error"