Type Systems

- Representation Independence

- Types as a safety measure in programming languages
  - Robin Milner: Well-typed programs don’t go wrong.

- Invariants and Equational properties

Example: Ariane 5

- Example: The Explosion of Ariane 5
- From the Official Report:
  - On 4 June 1996, the maiden flight of the Ariane 5 launcher ended in a failure. Only about 40 seconds after initiation of the flight sequence, at an altitude of about 3700 m, the launcher veered off its flight path, broke up and exploded.
  - The failure of the Ariane 501 was caused by the complete loss of guidance and attitude information 37 seconds after start of the main engine ignition sequence (30 seconds after lift-off). This loss of information was due to specification and design errors in the software of the inertial reference system.
  - The internal SRI software exception was caused during execution of a data conversion from 64-bit floating point to 16-bit signed integer value. The floating point number which was converted had a value greater than what could be represented by a 16-bit signed integer.

http://www.ima.umn.edu/~arnold/disasters/ariane5rep.html

Example: Ariane 5

- The Inquiry Board recommends:
  - R5 Review all flight software (including embedded software), and in particular:
    - identify all implicit assumptions made by the code and its justification documents on the values of quantities provided by the equipment, ...
  - R9 Include external (to the project) participants when reviewing specifications, code and justification documents. Make sure that these reviews consider the substance of arguments, rather than check that verifications have been made.
Type Systems

- Preventing execution errors by capturing them at compile time
  - Preventing a specific class of execution errors
  - Idea: Static analysis is easier than observing runtime behavior

- Safety
  - But... many languages are compiled to low-level unsafe code
  - Poses the challenge of compiler correctness

Typed Arithmetic Expressions

- Terms can evaluate to:
  - \( v ::= \)
    - \( \text{true} \)
    - \( \text{false} \)
    - \( \text{nv} \)
  - \( \text{nv} ::= \)
    - \( \text{0} \)
    - \( \text{succ \ nv} \)
  - Or...

Typed Arithmetic Expressions

- \( t ::= \)
  - \( \text{true} \)
  - \( \text{false} \)
  - \( \text{if \ t \ then \ t \ else \ t} \)
  - \( \text{\(0\)} \)
  - \( \text{\(\text{succ \ t}\)} \)
  - \( \text{\(\text{pred \ t}\)} \)
  - \( \text{\(\text{iszero \ t}\)} \)

Evaluation Rules

- \( \text{\(\text{if \ true \ then \ t_2 \ else \ t_3 \rightarrow t_2}\)} \) (E-IfTrue)
- \( \text{\(\text{if \ false \ then \ t_2 \ else \ t_3 \rightarrow t_3}\)} \) (E-IfFalse)
- \( \text{\(\text{if \ t_1 \ then \ t_2 \ else \ t_3 \rightarrow t_1\rightarrow t_3}\)} \) (E-If)
### Evaluation Rules

- \( t_1 \rightarrow t'_1 \) \hspace{1cm} (E-Succ)  \\
- \( \text{succ } t_1 \rightarrow \text{succ } t'_1 \) \hspace{1cm} (E-PredZero)  \\
- \( \text{pred } 0 \rightarrow 0 \) \hspace{1cm} (E-PredSucc)  \\
- \( \text{pred } \text{succ}(n v_1) \rightarrow n v_1 \) \hspace{1cm} (E-Pred)  \\
- \( t_1 \rightarrow t'_1 \) \hspace{1cm} (E-IsZeroZero)  \\
- \( \text{iszero } 0 \rightarrow \text{true} \) \hspace{1cm} (E-IsZeroSucc)  \\
- \( \text{iszero succ}(n v_1) \rightarrow \text{false} \) \hspace{1cm} (E-IsZero)  \\
- \( t_1 \rightarrow t'_1 \) \hspace{1cm} (E-IsZero)  \\
- \( \text{iszero } t_1 \rightarrow \text{iszero } t'_1 \)  

### Example

- \( \text{pred if true then iszero succ } 0 \text{ else } 0 \)  
- \( \text{pred(if (true) then (iszero (succ(0))) else (0))} \)  
- \( \text{pred } \text{(iszero succ (0))} \)  
- \( \text{pred false} \)

### Stuck Term

- No further evaluation rule applies that would bring the term into a normal form  
- Erroneous program

### A Type System can tell us without actual evaluation that a term does not get stuck

### Introducing a Type System

- Values can be of two different kinds in this term language  
- Boolean (Bool)  
  - E.g., the result of iszero  
  - The first argument of a conditional  
- Natural Numbers (Nat)  
  - E.g., the result of succ and pred  
  - E.g., the argument of iszero

- We can now say “a term has type T” when it obviously evaluates to a value of the appropriate form  
  - Obviously = statically, without performing evaluation  
- if true then false else true has type Bool  
- \( \text{pred(succ(succ(0)))} \) has type Nat

### Typing Rules

- \( \text{true} : \text{Bool} \) \hspace{1cm} (T-True)  
- \( \text{false} : \text{Bool} \) \hspace{1cm} (T-False)  
- \( t_1 : \text{Bool} \quad t_2 : T \quad t_3 : T \) \hspace{1cm} (T-If)
Typing Rules

- $0 : \text{Nat}$ (T-Zero)
- $t_1 : \text{Nat}$
  - $\text{succ} t_1 : \text{Nat}$ (T-Succ)
- $t_1 : \text{Nat}$
  - $\text{pred} t_1 : \text{Nat}$ (T-Pred)
- $t_1 : \text{Nat}$
  - $\text{iszero} t_1 : \text{Bool}$ (T-IsZero)

Typing Relation

- Def: The **Typing Relation** for arithmetic expressions is the smallest binary relation between terms and types satisfying all instances of the typing rules.
- Def: A term $t$ is **typeable** or **well-typed** if there is some $T$ such that $t : T$
- Idea: Answers of the sort: If the term has any type at all, then it must be type $\text{Bool}$.

Typing Derivation

- **if** $\text{iszero}$ $\theta$ then $\theta$ else $\text{succ}$ $\theta : \text{Nat}$

  \[
  \begin{align*}
  0 & : \text{Nat} \quad \text{(T-Zero)} \\
  \text{iszero} 0 & : \text{Nat} \quad \text{(T-IsZero)} \\
  0 & : \text{Nat} \quad \text{(T-Zero)} \\
  \text{succ} 0 & : \text{Nat} \quad \text{(T-Succ)} \\
  \end{align*}
  \]

  if $\text{iszero} 0$ then $0$ else $\text{succ} 0 : \text{Nat}$ (T-If)

Every pair $(t, T)$ in the typing relation can be justified by such a derivation tree.

Inversion Lemma

- **if** $\text{true} : R$, then $R = \text{Bool}$.
- **if** $\text{false} : R$, then $R = \text{Bool}$.
- **if** $t_3$ then $t_2$ else $t_3 : R$, then $t_1 : \text{Bool}$, $t_2 : R$, and $t_3 : R$.
- **if** $0 : R$, then $R = \text{Nat}$.
- **if** $\text{succ} t_1 : R$, then $R = \text{Nat}$ and $t_1 : \text{Nat}$.
- **if** $\text{pred} t_1 : R$, then $R = \text{Nat}$ and $t_1 : \text{Nat}$.
- **if** $\text{iszero} t_1 : R$, then $R = \text{Bool}$ and $t_1 : \text{Nat}$.

Practically describes an algorithm to construct a typing
Limitations of Typing
- What is the type of the following term:
  
  \[
  \text{if } \text{false} \text{ then } \text{true} \text{ else } 0
  \]

- Well-formed versus well-typed.
- Type systems are often conservative. They do not predict the outcome of every possible program but only cover cases in which a safe approximation can be made.

Safety: Soundness of Type Systems

- Progress
  - A well-typed term \( t \) is not stuck
    - If \( t : T \) then either \( t \) is a value or \( t \to t' \) for some \( t' \).

- Preservation
  - Types are preserved by one-step evaluation
    - If \( t : T \) and \( t \to t' \) then \( t' : T \)

Canonical Forms

- Lemma:
  - If \( v \) is a value of type \( \text{Bool} \), then \( v \) is either \text{true} or \text{false}.
  - If \( v \) is a value of type \( \text{Nat} \), then \( v \) is a numeric value.

- Progress:
  - Progress: If \( t : T \) then either \( t \) is a value or \( t \to t' \) for some \( t' \).

  - Proof by induction on a derivation of \( t : T \):
    - Case T-True: \( t = \text{true} \Rightarrow t \) is a value
    - Case T-False: \( t = \text{false} \Rightarrow t \) is a value
    - Case T-Zero: \( t = 0 \Rightarrow t \) is a value
Progress

Claim: If \( t : T \) then either \( t \) is a value or \( t \rightarrow t' \) for some \( t' \).

- **Case T-If:** Inversion lemma tells us that for \( t = \text{if } t_1 \text{ then } t_2 \text{ else } t_3 \),\
  \( t_1 : \text{Bool}, t_2 : T, t_3 : T \)

  - The following evaluation rules apply: E-IfTrue, E-IfFalse, and E-If

  Assume that the theorem holds for \( t_1 \).
  - If \( t_1 \) is a value, then \( t_2 \) and \( t_3 \) must be either true or false. In this case, the reduction can continue through E-IfTrue and E-IfFalse.
  - There exists a \( t_1' \) such that \( t_1 \rightarrow t_1' \) (since we know \( t_1 : \text{Bool} \)) and the reduction of \( t \) can be accomplished via E-If.

Preservation

Claim: If \( t : T \) and \( t \rightarrow t' \) then \( t' : T \).

- **Case T-True:** \( t = \text{true} \), \( T = \text{Bool} \), \( t \rightarrow t' \) is not possible.
- **Case T-False:** \( t = \text{true} \), \( T = \text{Bool} \), \( t \rightarrow t' \) is not possible.
- **Case T-Zero:** \( t = 0, T = \text{Nat} \), \( t \rightarrow t' \) is not possible.
- **Case T-Succ:** \( t = \text{succ } t_4, T = \text{Nat}, t_4 : \text{Nat} \)
  - Only E-Succ can derive \( a \rightarrow t' \), since \( t_4 : \text{Nat} \) by induction hypothesis

- **Case T-If:** \( t = \text{if } t_2 \text{ then } t_2 \text{ else } t_3, t_4 : \text{ Bool}, t_2 : T, t_3 : T \)
  - E-IfTrue: \( t_1 = \text{true}, t' = t_2, \) since \( t_2 : T \) also \( t' : T \)
  - E-IfFalse: symmetric with \( t' = t_3 \)
  - E-If: \( t' = \text{if } t_2' \text{ then } t_2 \text{ else } t_3, \) since \( t_4 : \text{ Bool} \) also \( t_2' : \text{Bool} \) and since \( t_2 : T \) and \( t_3 : T \) we can conclude that \( t' : T \).
- **Case T-Pred:** \( t = \text{pred } t_1, E-\text{PredZero} \), E-\text{PredSucc}, and E-Pred must be considered
- **Case T-IsZero:** \( t = \text{iszero } t_1, E-\text{IsZeroZero} \), E-\text{IsZeroSucc}, and E-\text{IsZero} must be considered.