Unification

Madhavan Mukund, **S P Suresh**

Programming Language Concepts

Lecture 25, 23 April 2024

Unification

- Start with a system of equations over **terms**

Unification

- Start with a system of equations over **terms**
- Find a **substitution** for variables that satisfies equations

Unification

- Start with a system of equations over **terms**
- Find a **substitution** for variables that satisfies equations
- Least constrained solution – **most general unifier (mgu)**

Terms

- Fix a set of function symbols and constants – **signature**

Terms

- Fix a set of function symbols and constants – **signature**
  - Each function symbol has an **arity**

Terms

- Fix a set of function symbols and constants – **signature**
  - Each function symbol has an **arity**
  - Constants are functions with arity 0

Terms

- Fix a set of function symbols and constants – **signature**
  - Each function symbol has an **arity**
  - Constants are functions with arity 0
- Terms are well-formed expressions, including **variables**

Terms

- Fix a set of function symbols and constants – **signature**
  - Each function symbol has an **arity**
  - Constants are functions with arity 0
- Terms are well-formed expressions, including **variables**
  - Every variable is a term

Terms

- Fix a set of function symbols and constants – **signature**
  - Each function symbol has an **arity**
  - Constants are functions with arity 0
- Terms are well-formed expressions, including **variables**
  - Every variable is a term
  - If f is a k-ary function symbols and $t_1, \ldots, t_k$ are terms, $f(t_1, \ldots, t_k)$ is also a term

Terms

- Fix a set of function symbols and constants – **signature**
  - Each function symbol has an **arity**
  - Constants are functions with arity 0
- Terms are well-formed expressions, including **variables**
  - Every variable is a term
  - If f is a k-ary function symbols and $t_1, \ldots, t_k$ are terms, $f(t_1, \ldots, t_k)$ is also a term
- **Example** – Types in λ-calculus

Terms

- Fix a set of function symbols and constants – **signature**
    - Each function symbol has an **arity**
    - Constants are functions with arity 0
- Terms are well-formed expressions, including **variables**
    - Every variable is a term
    - If f is a k-ary function symbols and $t_1, \ldots, t_k$ are terms, $f(t_1, \ldots, t_k)$ is also a term
- **Example** – Types in λ-calculus
    - *Int*, *Bool* …are constants

Terms

- Fix a set of function symbols and constants – **signature**
  - Each function symbol has an **arity**
  - Constants are functions with arity 0
- Terms are well-formed expressions, including **variables**
  - Every variable is a term
  - If f is a k-ary function symbols and $t_1, \ldots, t_k$ are terms, $f(t_1, \ldots, t_k)$ is also a term
- **Example** – Types in λ-calculus
  - $Int, Bool$ …are constants
  - $\rightarrow$, [], $(\cdot, \cdot)$ …are function symbols

Terms

- Fix a set of function symbols and constants – **signature**
  - Each function symbol has an **arity**
  - Constants are functions with arity 0
- Terms are well-formed expressions, including **variables**
  - Every variable is a term
  - If f is a k-ary function symbols and $t_1, \ldots, t_k$ are terms, $f(t_1, \ldots, t_k)$ is also a term
- **Example** – Types in λ-calculus
  - $Int, Bool$ …are constants
  - $\rightarrow$, [], $(\cdot, \cdot)$ …are function symbols
- **Notation**

Terms

- Fix a set of function symbols and constants – **signature**
  - Each function symbol has an **arity**
  - Constants are functions with arity 0
- Terms are well-formed expressions, including **variables**
  - Every variable is a term
  - If f is a k-ary function symbols and $t_1, \ldots, t_k$ are terms, $f(t_1, \ldots, t_k)$ is also a term
- **Example** – Types in λ-calculus
  - $Int, Bool$ …are constants
  - $\rightarrow$, $[]$, $(\cdot, \cdot)$ …are function symbols
- **Notation**
  - $a, b, c, \ldots, f, g, \ldots, x, y, \ldots$ are function symbols

Terms

- Fix a set of function symbols and constants – **signature**
  - Each function symbol has an **arity**
  - Constants are functions with arity 0
- Terms are well-formed expressions, including **variables**
  - Every variable is a term
  - If f is a k-ary function symbols and $t_1, \ldots, t_k$ are terms, $f(t_1, \ldots, t_k)$ is also a term
- **Example** – Types in λ-calculus
  - $Int, Bool$ …are constants
  - $\rightarrow$, [], $(\cdot, \cdot)$ …are function symbols
- **Notation**
  - $a, b, c, \ldots, f, g, \ldots, x, y, \ldots$ are function symbols
  - $A, B, C, F, X, Y, \ldots$ are variables

Unification

$$f(X) = f(f(a))$$
$$g(Y) = g(Z)$$

- **Substitution**: assigns a term to each variable X, Y, Z

Unification

$$f(X) = f(f(a))$$
$$g(Y) = g(Z)$$

- **Substitution**: assigns a term to each variable X, Y, Z
- **Unifier**: substitution that satisfies equations

Unification

$$f(X) = f(f(a))$$
$$g(Y) = g(Z)$$

- **Substitution**: assigns a term to each variable X, Y, Z
- **Unifier**: substitution that satisfies equations
- For instance, $\{X := f(a), Y := g(a), Z := g(a)\} = \theta$

Unification

$$f(X) = f(f(a))$$
$$g(Y) = g(Z)$$

- **Substitution**: assigns a term to each variable X, Y, Z
- **Unifier**: substitution that satisfies equations
- For instance, $\{X := f(a), Y := g(a), Z := g(a)\} = \theta$
- $t\theta$: apply substitution θ to term t (not $\theta(t)$!)

Unification

$$f(X) = f(f(a))$$
$$g(Y) = g(Z)$$

- **Substitution**: assigns a term to each variable X, Y, Z
- **Unifier**: substitution that satisfies equations
- For instance, $\{X := f(a), Y := g(a), Z := g(a)\} = \theta$
- $t\theta$: apply substitution θ to term t (not $\theta(t)$!)
  - Apply substitution in parallel

Unification

$$f(X) = f(f(a))$$

$$g(Y) = g(Z)$$

- **Substitution**: assigns a term to each variable X, Y, Z
- **Unifier**: substitution that satisfies equations
- For instance, $\{X := f(a), Y := g(a), Z := g(a)\} = \theta$
- $t\theta$: apply substitution θ to term t (not $\theta(t)$!)
  - Apply substitution in parallel
  - $t = g(p(X), q(f(Y)))$

Unification

$$f(X) = f(f(a))$$

$$g(Y) = g(Z)$$

- **Substitution**: assigns a term to each variable X, Y, Z
- **Unifier**: substitution that satisfies equations
- For instance, $\{X := f(a), Y := g(a), Z := g(a)\} = \theta$
- $t\theta$: apply substitution θ to term t (not $\theta(t)$!)
  - Apply substitution in parallel
  - $t = g(p(X), q(f(Y)))$
  - $\gamma = \{X := Y, Y := f(a)\}$

Unification

$$f(X) = f(f(a))$$

$$g(Y) = g(Z)$$

- **Substitution**: assigns a term to each variable X, Y, Z
- **Unifier**: substitution that satisfies equations
- For instance, $\{X := f(a), Y := g(a), Z := g(a)\} = \theta$
- $t\theta$: apply substitution θ to term t (not $\theta(t)$!)
  - Apply substitution in parallel
  - $t = g(p(X), q(f(Y)))$
  - $\gamma = \{X := Y, Y := f(a)\}$
  - $t\gamma = g(p(Y), q(f(f(a))))$ – $g(p(Y))$, **not** $g(p(f(a)))$

Unification

$$f(X) = f(f(a))$$
$$g(Y) = g(Z)$$

- Many solutions are possible

Unification

$$f(X) = f(f(a))$$
$$g(Y) = g(Z)$$

- Many solutions are possible
  - $\theta = \{X := f(a), Y := g(a), Z := g(a)\}$

Unification

$$f(X) = f(f(a))$$

$$g(Y) = g(Z)$$

- Many solutions are possible
  - $\theta = \{X := f(a), Y := g(a), Z := g(a)\}$
  - $\theta' = \{X := f(a), Y := a, Z := a\}$

Unification

$$f(X) = f(f(a))$$
$$g(Y) = g(Z)$$

- Many solutions are possible
  - $\theta = \{X := f(a), Y := g(a), Z := g(a)\}$
  - $\theta' = \{X := f(a), Y := a, Z := a\}$
  - $\theta'' = \{X := f(a), Y := Z\}$

Unification

$$f(X) = f(f(a))$$

$$g(Y) = g(Z)$$

- Many solutions are possible
  - $\theta = \{X := f(a), Y := g(a), Z := g(a)\}$
  - $\theta' = \{X := f(a), Y := a, Z := a\}$
  - $\theta'' = \{X := f(a), Y := Z\}$
- θ'' is the **least constrained**

Unification

$$f(X) = f(f(a))$$

$$g(Y) = g(Z)$$

- Many solutions are possible
  - $\theta = \{X := f(a), Y := g(a), Z := g(a)\}$
  - $\theta' = \{X := f(a), Y := a, Z := a\}$
  - $\theta'' = \{X := f(a), Y := Z\}$
- θ'' is the **least constrained**
- Any solution γ can be broken up into θ'' followed by another substitution

Unification

$$f(X) = f(f(a))$$
$$g(Y) = g(Z)$$

- Many solutions are possible
  - $\theta = \{X := f(a), Y := g(a), Z := g(a)\}$
  - $\theta' = \{X := f(a), Y := a, Z := a\}$
  - $\theta'' = \{X := f(a), Y := Z\}$
- θ'' is the **least constrained**
- Any solution γ can be broken up into θ'' followed by another substitution
  - θ is θ'' followed by $\{Y := g(a)\}$

Unification

$$f(X) = f(f(a))$$

$$g(Y) = g(Z)$$

- Many solutions are possible
  - $\theta = \{X := f(a), Y := g(a), Z := g(a)\}$
  - $\theta' = \{X := f(a), Y := a, Z := a\}$
  - $\theta'' = \{X := f(a), Y := Z\}$
- θ'' is the **least constrained**
- Any solution γ can be broken up into θ'' followed by another substitution
  - θ is θ'' followed by $\{Y := g(a)\}$
- Least constrained solution – **most general unifier**

Obstacles to unification

- Equations of the form $p(\cdots) = q(\cdots)$

Obstacles to unification

- Equations of the form $p(\cdots) = q(\cdots)$
    - Outermost function symbols do not agree

Obstacles to unification

- Equations of the form $p(\cdots) = q(\cdots)$
  - Outermost function symbols do not agree
  - No substitution can equate the two sides

Obstacles to unification

- Equations of the form $p(\cdots) = q(\cdots)$
  - Outermost function symbols do not agree
  - No substitution can equate the two sides
  - **Example**: No substitution for type variables can equate $[p]$ and (q, r)

Obstacles to unification

- Equations of the form $p(\cdots) = q(\cdots)$
  - Outermost function symbols do not agree
  - No substitution can equate the two sides
  - **Example**: No substitution for type variables can equate $[p]$ and (q, r)
- Equations of the form $X = f(\cdots X \cdots)$

Obstacles to unification

- Equations of the form $p(\cdots) = q(\cdots)$
  - Outermost function symbols do not agree
  - No substitution can equate the two sides
  - **Example**: No substitution for type variables can equate $[p]$ and (q, r)
- Equations of the form $X = f(\cdots X \cdots)$
  - Any substitution for X applies also to the X nested inside f

Obstacles to unification

- Equations of the form $p(\cdots) = q(\cdots)$
  - Outermost function symbols do not agree
  - No substitution can equate the two sides
  - **Example**: No substitution for type variables can equate $[p]$ and (q, r)
- Equations of the form $X = f(\cdots X \cdots)$
  - Any substitution for X applies also to the X nested inside f
- These are the **only two** obstacles to unification

A unification algorithm

- Start with system of equations

$$\begin{aligned} l_1 &= r_1 \\ l_2 &= r_2 \\ &\vdots \\ l_n &= r_n \end{aligned}$$

A unification algorithm

- Start with system of equations

$$\begin{aligned} l_1 &= r_1 \\ l_2 &= r_2 \\ &\vdots \\ l_n &= r_n \end{aligned}$$

- Perform a sequence of transformations on these equations till no more transformations apply

Unification algorithm – transformations

1. $t = X$, t not a variable $\rightsquigarrow$ $X = t$

Unification algorithm – transformations

1. $t = X$, t not a variable $\rightsquigarrow$ $X = t$
2. Erase equations of the form $X = X$

Unification algorithm – transformations

1. $t = X$, t not a variable $\quad\rightsquigarrow\quad$ $X = t$
2. Erase equations of the form $X = X$
3. Let $t = t'$ where t is $f(\cdots)$ and t' is $f'(\cdots)$

Unification algorithm – transformations

1. $t = X$, t not a variable $\rightsquigarrow$ $X = t$
2. Erase equations of the form $X = X$
3. Let $t = t'$ where t is $f(\cdots)$ and t' is $f'(\cdots)$
   - $f \neq f'$ $\rightsquigarrow$ terminate: unification not possible

Unification algorithm – transformations

1. $t = X$, t not a variable $\rightsquigarrow$ $X = t$
2. Erase equations of the form $X = X$
3. Let $t = t'$ where t is $f(\cdots)$ and t' is $f'(\cdots)$
   - $f \neq f'$ $\rightsquigarrow$ terminate: unification not possible
   - $f(t_1, \ldots, t_k) = f(t'_1, \ldots, t'_k)$

Unification algorithm – transformations

1. $t = X$, t not a variable $\quad\rightsquigarrow\quad X = t$
2. Erase equations of the form $X = X$
3. Let $t = t'$ where t is $f(\cdots)$ and t' is $f'(\cdots)$
   - $f \neq f' \quad\rightsquigarrow\quad$ terminate: unification not possible
   - $f(t_1, \ldots, t_k) = f(t'_1, \ldots, t'_k)$
   - Replace by k new equations $t_1 = t'_1, \ldots, t_k = t'_k$

Unification algorithm – transformations

1. $t = X$, t not a variable $\rightsquigarrow$ $X = t$
2. Erase equations of the form $X = X$
3. Let $t = t'$ where t is $f(\cdots)$ and t' is $f'(\cdots)$
    - $f \neq f'$ $\rightsquigarrow$ terminate: unification not possible
    - $f(t_1, \ldots, t_k) = f(t'_1, \ldots, t'_k)$
    - Replace by k new equations $t_1 = t'_1, \ldots, t_k = t'_k$
4. $X = t$, X is a proper subterm of t $\rightsquigarrow$ terminate: unification not possible

Unification algorithm – transformations

1. $t = X$, t not a variable $\rightsquigarrow$ $X = t$
2. Erase equations of the form $X = X$
3. Let $t = t'$ where t is $f(\cdots)$ and t' is $f'(\cdots)$
   - $f \neq f'$ $\rightsquigarrow$ terminate: unification not possible
   - $f(t_1, \ldots, t_k) = f(t'_1, \ldots, t'_k)$
   - Replace by k new equations $t_1 = t'_1, \ldots, t_k = t'_k$
4. $X = t$, X is a proper subterm of t $\rightsquigarrow$ terminate: unification not possible
5. $X = t$, X does not occur in t $\rightsquigarrow$ Replace all other occurrences of X by t

Unification algorithm – correctness

- The algorithm terminates

Unification algorithm – correctness

- The algorithm terminates
  - Rules 1–4 can be used only a finite number of times without using Rule 5

Unification algorithm – correctness

- The algorithm terminates
  - Rules 1–4 can be used only a finite number of times without using Rule 5
  - Rule 5 can be used at most once for each variable

Unification algorithm – correctness

- The algorithm terminates
  - Rules 1–4 can be used only a finite number of times without using Rule 5
  - Rule 5 can be used at most once for each variable
- When the algorithm terminates, all equations are of the form $X_i = t_i$

Unification algorithm – correctness

- The algorithm terminates
  - Rules 1–4 can be used only a finite number of times without using Rule 5
  - Rule 5 can be used at most once for each variable
- When the algorithm terminates, all equations are of the form $X_i = t_i$
  - exactly one occurrence of X_i

Unification algorithm – correctness

- The algorithm terminates
  - Rules 1–4 can be used only a finite number of times without using Rule 5
  - Rule 5 can be used at most once for each variable
- When the algorithm terminates, all equations are of the form $X_i = t_i$
  - exactly one occurrence of X_i
- Define a substitution $\{X_1 := t_1, \ldots, X_m := t_m\}$

Unification algorithm – correctness

- The algorithm terminates
  - Rules 1–4 can be used only a finite number of times without using Rule 5
  - Rule 5 can be used at most once for each variable
- When the algorithm terminates, all equations are of the form $X_i = t_i$
  - exactly one occurrence of X_i
- Define a substitution $\{X_1 := t_1, \ldots, X_m := t_m\}$
- This substitution is a unifier

Unification algorithm – correctness

- The algorithm terminates
  - Rules 1–4 can be used only a finite number of times without using Rule 5
  - Rule 5 can be used at most once for each variable
- When the algorithm terminates, all equations are of the form $X_i = t_i$
  - exactly one occurrence of X_i
- Define a substitution $\{X_1 := t_1, \ldots, X_m := t_m\}$
- This substitution is a unifier
  - Each of the transformations preserves solutions

Unification algorithm – correctness

- The algorithm terminates
  - Rules 1–4 can be used only a finite number of times without using Rule 5
  - Rule 5 can be used at most once for each variable
- When the algorithm terminates, all equations are of the form $X_i = t_i$
  - exactly one occurrence of X_i
- Define a substitution $\{X_1 := t_1, \ldots, X_m := t_m\}$
- This substitution is a unifier
  - Each of the transformations preserves solutions
- The substitution is also an mgu

Unification algorithm – example

$$g(Y) = X$$

$$f(X, h(X), Y) = f(g(Z), W, Z)$$

Unification algorithm – example

$$g(Y) = X$$
$$f(X, h(X), Y) = f(g(Z), W, Z)$$

$$X = g(Y)$$
$$f(X, h(X), Y) = f(g(Z), W, Z)$$

Unification algorithm – example

$$g(Y) = X$$
$$f(X, h(X), Y) = f(g(Z), W, Z)$$

$$X = g(Y)$$
$$X = g(Z)$$
$$h(X) = W$$
$$Y = Z$$

$$X = g(Y)$$
$$f(X, h(X), Y) = f(g(Z), W, Z)$$

Unification algorithm – example

$$g(Y) = X \qquad X = g(Y)$$

$$f(X, h(X), Y) = f(g(Z), W, Z) \qquad X = g(Z)$$

$$h(X) = W$$

$$Y = Z$$

$$X = g(Y)$$

$$f(X, h(X), Y) = f(g(Z), W, Z)$$

$$g(Z) = g(Y)$$

$$X = g(Z)$$

$$h(g(Z)) = W$$

$$Y = Z$$

Unification algorithm – example

$$g(Y) = X$$
$$f(X, h(X), Y) = f(g(Z), W, Z)$$

$$X = g(Y)$$
$$f(X, h(X), Y) = f(g(Z), W, Z)$$

$$g(Z) = g(Y)$$
$$X = g(Z)$$
$$h(g(Z)) = W$$
$$Y = Z$$

$$X = g(Y)$$
$$X = g(Z)$$
$$h(X) = W$$
$$Y = Z$$

$$Z = Y$$
$$X = g(Z)$$
$$h(g(Z)) = W$$
$$Y = Z$$

Unification algorithm – example

$$g(Y) = X$$
$$f(X, h(X), Y) = f(g(Z), W, Z)$$

$$X = g(Y)$$
$$X = g(Z)$$
$$h(X) = W$$
$$Y = Z$$

$$Z = Y$$
$$X = g(Z)$$
$$h(g(Z)) = W$$
$$Y = Z$$

$$X = g(Y)$$
$$f(X, h(X), Y) = f(g(Z), W, Z)$$

$$g(Z) = g(Y)$$
$$X = g(Z)$$
$$h(g(Z)) = W$$
$$Y = Z$$

$$Z = Z$$
$$X = g(Z)$$
$$h(g(Z)) = W$$
$$Y = Z$$

Unification algorithm – example

$$g(Y) = X$$
$$f(X, h(X), Y) = f(g(Z), W, Z)$$

$$X = g(Y)$$
$$X = g(Z)$$
$$h(X) = W$$
$$Y = Z$$

$$Z = Y$$
$$X = g(Z)$$
$$h(g(Z)) = W$$
$$Y = Z$$

$$X = g(Z)$$
$$h(g(Z)) = W$$
$$Y = Z$$

$$X = g(Y)$$
$$f(X, h(X), Y) = f(g(Z), W, Z)$$

$$g(Z) = g(Y)$$
$$X = g(Z)$$
$$h(g(Z)) = W$$
$$Y = Z$$

$$Z = Z$$
$$X = g(Z)$$
$$h(g(Z)) = W$$
$$Y = Z$$

Unification algorithm – example

$$g(Y) = X$$
$$f(X, h(X), Y) = f(g(Z), W, Z)$$

$$X = g(Y)$$
$$f(X, h(X), Y) = f(g(Z), W, Z)$$

$$X = g(Y)$$
$$X = g(Z)$$
$$h(X) = W$$
$$Y = Z$$

$$g(Z) = g(Y)$$
$$X = g(Z)$$
$$h(g(Z)) = W$$
$$Y = Z$$

$$Z = Y$$
$$X = g(Z)$$
$$h(g(Z)) = W$$
$$Y = Z$$

$$Z = Z$$
$$X = g(Z)$$
$$h(g(Z)) = W$$
$$Y = Z$$

$$X = g(Z)$$
$$h(g(Z)) = W$$
$$Y = Z$$

$$X = g(Z)$$
$$W = h(g(Z))$$
$$Y = Z$$

Unification algorithm – example

Equations: $g(Y) = X \qquad f(X, h(X), Y) = f(g(Z), W, Z)$

Unifier: $\{X := g(Z), Y := Z, W := h(g(Z))\}$