

# First-order Predicate Logic Normal Forms

# Substitution

- ▶ Substitutions replace *free* variables by terms.  
(They are mappings from variables to terms)
- ▶ By  $[t/x]$  we denote the substitution that replaces  $x$  by  $t$ .
- ▶ The notation  $F[t/x]$  (“ $F$  with  $t$  for  $x$ ”) denotes the result of replacing all *free* occurrences of  $x$  in  $F$  by  $t$ .

## Example

$$(\forall x P(x) \wedge Q(x))[f(y)/x] = \forall x P(x) \wedge Q(f(y))$$

- ▶ Similarly for substitutions in terms:  
 $u[t/x]$  is the result of replacing  $x$  by  $t$  in term  $u$ .

## Example

$$(f(x))[g(x)/x] = f(g(x))$$

## Substitution: Variations

Simultaneous substitution of  $t_i$  for  $x_i$  is denoted by  $[t_1/x_1, \dots, t_n/x_n]$ .

### Example

$$(P(x, y))[f(y)/x, b/y] = P(f(y), b)$$

Composition of substitutions is denoted by juxtaposition:  $[t_1/x][t_2/y]$  first substitutes  $t_1$  for  $x$  and then substitutes  $t_2$  for  $y$ .

### Example

$$(P(x, y))[f(y)/x][b/y] = (P(f(y), y))[b/y] = P(f(b), b)$$

# Variable capture

## Warning

If  $t$  contains a variable that is bound in  $F$ ,  
substitution may lead to **variable capture**:

$$(\forall x P(x, y))[f(x)/y] = \forall x P(x, f(x))$$

Variable capture should be avoided

# Substitution lemmas

## Lemma (Substitution Lemma)

*If  $t$  contains no variable bound in  $F$  then*

$$\mathcal{A}(F[t/x]) = (\mathcal{A}[\mathcal{A}(t)/x])(F)$$

**Proof** by structural induction on  $F$   
with the help of the corresponding lemma on terms:

## Lemma

$$\mathcal{A}(u[t/x]) = (\mathcal{A}[\mathcal{A}(t)/x])(u)$$

**Proof** by structural induction on  $u$

# Warning

The notation  $.[./.]$  is heavily overloaded:

## Substitution in syntactic objects

$F[G/A]$  in propositional logic

$F[t/x]$

$u[t/x]$  where  $u$  is a term

## Function update

$\mathcal{A}[v/A]$  where  $\mathcal{A}$  is a propositional assignment

$\mathcal{A}[d/x]$  where  $\mathcal{A}$  is a structure and  $d \in U_{\mathcal{A}}$

# Rectified Formulas

## Definition

A formula is **rectified** if no variable occurs both bound and free and if all quantifiers in the formula bind different variables.

## Lemma

Let  $F = QxG$  be a formula where  $Q \in \{\forall, \exists\}$ .

Let  $y$  be a variable that does not occur free in  $G$ .

Then  $F \equiv QyG[y/x]$ .

## Lemma

Every formula is equivalent to a rectified formula.

## Example

$$\forall x P(x, y) \wedge \exists x \exists y Q(x, y) \equiv \forall x' P(x', y) \wedge \exists x \exists y' Q(x, y')$$

# Prenex form

## Definition

A formula is in **prenex form** if it has the form

$$Q_1 y_1 Q_2 y_2 \dots Q_n y_n F$$

where  $Q_i \in \{\exists, \forall\}$ ,  $n \geq 0$ , and  $F$  is quantifier-free.



# Prenex form

## Theorem

Every formula is equivalent to a rectified formula in prenex form (a formula in **RPF**).

**Proof** First construct an equivalent rectified formula. Then pull the quantifiers to the front using the following equivalences from left to right as long as possible:

$$\neg \forall x F \equiv \exists x \neg F$$

$$\neg \exists x F \equiv \forall x \neg F$$

$$Qx F \wedge G \equiv Qx (F \wedge G)$$

$$F \wedge Qx G \equiv Qx (F \wedge G)$$

$$Qx F \vee G \equiv Qx (F \vee G)$$

$$F \vee Qx G \equiv Qx (F \vee G)$$

For the last four rules note that the formula is rectified!

## Skolem form

The **Skolem form** of a formula  $F$  in RPF is the result of applying the following algorithm to  $F$ :

**while**  $F$  contains an existential quantifier **do**

Let  $F = \forall y_1 \forall y_2 \dots \forall y_n \exists z G$

(the block of universal quantifiers may be empty)

Let  $f$  be a **fresh** function symbol of arity  $n$   
that does not occur in  $F$ .

$F := \forall y_1 \forall y_2 \dots \forall y_k G[f(y_1, y_2, \dots, y_n)/z]$

i.e. remove the outermost existential quantifier in  $F$  and  
replace every occurrence of  $z$  in  $G$  by  $f(y_1, y_2, \dots, y_n)$

### Example

$\exists x \forall y \exists z \forall u \exists v P(x, y, z, u, v)$

### Theorem

*A formula in RPF and its Skolem form are equisatisfiable.*

## Summary: conversion to Skolem form

**Input:** a formula  $F$

**Output:** an equisatisfiable, rectified, closed formula  
in Skolem form  $\forall y_1 \dots \forall y_k G$  where  $G$  is quantifier-free

1. Rectify  $F$  by systematic renaming of bound variables.  
The result is a formula  $F_1$  equivalent to  $F$ .
2. Let  $y_1, y_2, \dots, y_n$  be the variables occurring free in  $F_1$ .  
Produce the formula  $F_2 = \exists y_1 \exists y_2 \dots \exists y_n F_1$ .  
 $F_2$  is equisatisfiable with  $F_1$ , rectified and closed.
3. Produce a formula  $F_3$  in RPF equivalent to  $F_2$ .
4. Eliminate the existential quantifiers in  $F_3$   
by transforming  $F_3$  into its Skolem form  $F_4$ .  
The formula  $F_4$  is equisatisfiable with  $F_3$ .

## Exercise

Which formulas are rectified, in prenex, or Skolem form?

	R	P	S
$\forall x(T(x) \vee C(x) \vee D(x))$			
$\exists x \exists y(C(y) \vee B(x, y))$			
$\neg \exists x C(x) \leftrightarrow \forall x \neg C(x)$			
$\forall x(C(x) \rightarrow S(x)) \rightarrow \forall y(\neg C(y) \rightarrow \neg S(y))$			