

# THE LANGUAGE OF FIRST-ORDER LOGIC AND KNOWLEDGE REPRESENTATION

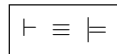
In the preceding example, we used the language of standard First-Order Logic (FOL) for the formal representation of knowledge and reasoning.

*Now:* Take a closer look at the language of FOL

## Remember:

*Operations* can be defined formally (**proof theory**)

and receive semantic properties (**truth theory**):



## Contents:

- A brief summary of the classical introduction to propositional and quantificational ("predicate") logic,
- Another example,
- Reasoning problems, decidability, and expressive power,
- Extensions motivated by requirements for knowledge representation and automated reasoning.

**Remark:** For a pragmatically based, constructive foundation of logic by means of dialogue games

$\Rightarrow$  *Excursus:* "A Constructive Introduction to First Order Logic"

## Calculi of Formal Logic

(Summary of the Classical Approach)

Distinction: **Provability** and **Truth**

*Formal theory*  $\mathcal{T}$ :

1. Symbols, expressions (sequences of symbols)
2. Well formed formulas (wff)  $\subset$  expressions
3. Axioms  $\subset$  wffs
4. Inference rules (relations on wffs)

## Important concepts:

- *Schema:* statement form;  
*Instance:* wff obtained by substitution.  
Infinite axiom systems via axiom schemata, metavariables
- $P$  is *deducible* from  $\Sigma$  in  $\mathcal{T}$ :  $\Sigma \vdash_{\mathcal{T}} P$  ;  
Concepts: *derivation, proof*  
 $P$  theorem:  $\vdash_{\mathcal{T}} P$ ,  $P$  provable in  $\mathcal{T}$
- *Interpretations:* meaning for each symbol s.t. any wff can be understood as a true or false statement in the interpretation.
- *Model:* interpretation for a set of wffs s.t. every wff is true in interpretation.  
Model for a theory  $\mathcal{T} \Leftrightarrow$  model for the set of theorems of  $\mathcal{T}$

- *Completeness of  $\mathcal{T}$* : Every sentence that is true in all interpretations is provable in  $\mathcal{T}$  (truth  $\rightarrow$  provability)
- *Soundness of  $\mathcal{T}$* : Every provable sentence is true in all interpretations (provability  $\rightarrow$  truth)
- *Decidability of  $\mathcal{T}$* : Effective procedure exists that will determine provability for any sentence
- *Consistency of  $\mathcal{T}$* : contains no wff s.t. wff and its negation are provable

## Fundamentals of Propositional Logic: Syntax

$\Rightarrow$  Logic course (e.g. R. Davis: Truth, Deduction and Computation, 1989)

### Syntax:

- Symbols
  - Proposition letters (variables), subscripted
  - Connectives:  $\rightarrow, \neg$
  - Auxiliary symbols:  $(, )$
- Well-formed formulas (wffs)
  - Prime (atomic) formulas: proposition letters
  - Compound formulas: If  $P, Q$  are wffs, so  $\neg P$  and  $P \rightarrow Q$

- Further propositional connectives (redundant):
  - $A \wedge B \Leftrightarrow \neg(A \rightarrow \neg B)$
  - $A \vee B \Leftrightarrow \neg A \rightarrow B$
  - $A \leftrightarrow B \Leftrightarrow (A \rightarrow B) \wedge (B \rightarrow A)$

## Fundamentals of Propositional Logic: Semantics

### Semantics: Truth

Truth tables

$A$	$\neg A$
T	⊥
⊥	T

$A$	$B$	$A \rightarrow B$
T	T	T
T	⊥	⊥
⊥	T	T
⊥	⊥	T

$A$	$B$	$A \vee B$
T	T	T
T	⊥	T
⊥	T	T
⊥	⊥	⊥

$A$	$B$	$A \wedge B$
T	T	T
T	⊥	⊥
⊥	T	⊥
⊥	⊥	⊥

$A$	$B$	$A \leftrightarrow B$
T	T	T
T	⊥	⊥
⊥	T	⊥
⊥	⊥	T

## Fundamentals of Propositional Logic: Semantics (2)

- Truth tables: all interpretations
- Connectives as logical functions
- $P$  *logically valid*: true in all interpretations  
 $\models P$  (tautology)
- *Satisfiability*: wff is true in (at least) one interpretation
- *Unsatisfiability* of  $P$ :  $\neg P$  is logically valid (a contradiction)

## Propositional Logic: Deduction

**Proof theory:** Deduction

Axiomatisation in Hilbert style:

*Axioms of  $\mathcal{L}$*

$\mathcal{L}1$   $(A \rightarrow (B \rightarrow A))$

$\mathcal{L}2$   $((A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (B \rightarrow C)))$

$\mathcal{L}3$   $((\neg B) \rightarrow (\neg A)) \rightarrow (((\neg B) \rightarrow A) \rightarrow B)$

*Inference rule:* Modus Ponens

$MP$   $A, A \rightarrow B \vdash B$

## Propositional Logic: Proof Example

*Lemma:*  $A \rightarrow A$

*Proof:*

1.  $(A \rightarrow ((A \rightarrow A) \rightarrow A)) \rightarrow ((A \rightarrow (A \rightarrow A)) \rightarrow (A \rightarrow A))$   
an instance of axiom schema  $\mathcal{L}2$ , with  $A$  for  $A$ ,  $(A \rightarrow A)$  for  $B$ ,  $A$  for  $C$
2.  $A \rightarrow ((A \rightarrow A) \rightarrow A)$   
axiom schema  $\mathcal{L}1$  with  $A$  for  $A$ ,  $(A \rightarrow A)$  for  $B$
3.  $(A \rightarrow (A \rightarrow A)) \rightarrow (A \rightarrow A)$  by MP on 1 and 2
4.  $A \rightarrow (A \rightarrow A)$  axiom schema  $\mathcal{L}1$ ,  $A$  for  $A$ ,  $A$  for  $B$
5.  $A \rightarrow A$  by MP on 3 and 4

## Propositional Logic: Metatheory

- Deduction theorem:  
Let  $\Sigma$  set of wffs,  $A, B$  wffs, and  $\Sigma, A \vdash B$ , then  $\Sigma \vdash A \rightarrow B$ .  
For  $\Sigma = \emptyset$ : If  $A \vdash B$  then  $\vdash A \rightarrow B$ .
- Completeness of  $\mathcal{L}$ : If  $\models P$  then  $\vdash P$
- Soundness of  $\mathcal{L}$ : If  $\vdash P$  then  $\models P$
- Hence, truth  $\equiv$  deduction, i.e.  $\vdash P \Leftrightarrow \models P$
- Consistency
- Decidability

## Seven Derived Inference Rules for Propositional Logic Natural Deduction Style

- Modus Ponens or subjunction elimination

$$\frac{\alpha \rightarrow \beta, \alpha}{\beta}$$

- And elimination

$$\frac{\alpha_1 \wedge \alpha_2 \wedge \dots \wedge \alpha_n}{\alpha_i}$$

- And introduction

$$\frac{\alpha_1, \alpha_2, \dots, \alpha_n}{\alpha_1 \wedge \alpha_2 \wedge \dots \wedge \alpha_n}$$

Complementary literal elimination (ground resolution) provides the computational basis for propositional logic.

- Or introduction

$$\frac{\alpha_i}{\alpha_1 \vee \alpha_2 \vee \dots \vee \alpha_n}$$

- Elimination of double negation

$$\frac{\neg\neg\alpha}{\alpha}$$

- Unit RESOLUTION

$$\frac{\alpha \vee \beta, \neg\beta}{\alpha}$$

- RESOLUTION

$$\frac{\alpha \vee \beta, \neg\beta \vee \gamma}{\alpha \vee \gamma} \text{ or equivalently } \frac{\neg\alpha \rightarrow \beta, \beta \rightarrow \gamma}{\neg\alpha \rightarrow \gamma}$$

## Soundness of the Resolution Inference Rule: Truth Table

$\alpha$	$\beta$	$\gamma$	$\alpha \vee \beta$	$\neg\beta \vee \gamma$	$\alpha \vee \gamma$
<i>False</i>	<i>False</i>	<i>False</i>	<i>False</i>	<i>True</i>	<i>False</i>
<i>False</i>	<i>False</i>	<i>True</i>	<i>False</i>	<i>True</i>	<i>True</i>
<i>False</i>	<i>True</i>	<i>False</i>	<i>True</i>	<i>False</i>	<i>False</i>
<i>False</i>	<i>True</i>	<i>True</i>	<i>True</i>	<i>True</i>	<i>True</i>
<i>True</i>	<i>False</i>	<i>False</i>	<i>True</i>	<i>True</i>	<i>True</i>
<i>True</i>	<i>False</i>	<i>True</i>	<i>True</i>	<i>True</i>	<i>True</i>
<i>True</i>	<i>True</i>	<i>False</i>	<i>True</i>	<i>False</i>	<i>True</i>
<i>True</i>	<i>True</i>	<i>True</i>	<i>True</i>	<i>True</i>	<i>True</i>

## Fundamentals of Quantificational Logic ("Predicate Calculus")

Investigation of truth and falseness of composed formulas whose atomic parts are no longer certain sentences, but *sentence forms* (parameterized sentences).

Introduction of *variables* and *instantiation* (substitution);  
Sentence forms determine relations expressed by *predicates* over their arguments.

### First order theories:

Restriction of arguments to terms which can be constructed from constants, variables and functional expressions and quantification over (simple) variables.

## Fundamentals of Quantificational Logic: Syntax

### Symbols of first order language:

- Connectives:  $\neg, \rightarrow, \wedge$
- Auxiliary symbols:  $(, ), ", "$
- Variables:  $x, x_1, x_2, \dots$
- Constant symbols:  $a, a_1, a_2, \dots$
- Function symbols:  $f_k^n$  ( $n, k \in \mathbf{N}$ )
- Predicate symbols:  $A_k^n$  ( $n, k \in \mathbf{N}$ )

### Well-formed formulas:

- term ::= variable | constant | function\_letter "(" termlist ")"
- termlist ::= term | term ", " termlist
- prime\_formula ::= predicate\_letter "(" termlist ")"
- wff ::= prime\_formula | "(" ¬ " wff ")"  
| "(" " wff " → " wff ")" | "(" ( ∧ " variable ")" wff ")"

**Existential quantifier:**  $((\forall x)A) \Leftrightarrow \neg((\exists x)\neg A)$

**Scope** of a quantifier: wff to which it applies

**Bound** and **free** variables

## Fundamentals of Quantificational Logic: Semantics

### Interpretation:

1. Domain  $\mathcal{D} \neq \emptyset$
2. Assignment to each  $n$ -ary predicate symbol  $A_k^n$  of an  $n$ -place relation in  $\mathcal{D}$
3. Assignment to each  $n$ -ary function symbol  $f_k^n$  of an  $n$ -place operation closed over  $\mathcal{D}$ :  $\mathcal{D}^n \rightarrow \mathcal{D}$
4. Assignment to each individual constant  $a_i$  of some fixed element of  $\mathcal{D}$

**Closed** wff  $\Leftrightarrow$  wff containing no free variables

## Fundamentals of Quantificational Logic: Semantics (2)

- Satisfiability in an interpretation, defined recursively;  
e.g. A sequence of elements  $s$  of  $\mathcal{D}$  satisfies  $\neg A$  if and only if  $s$  does not satisfy  $A$ .  
Note: Circularity on metalevel!
- Truth: A wff is true (for a given interpretation) iff every sequence in the set of sequences  $\Sigma$  satisfies  $A$ .  $A$  is false iff no sequence in  $\Sigma$  satisfies  $A$
- Logical validity of a wff  $A$  iff  $A$  is true for every interpretation.
- Satisfiability of a wff  $A$  iff there is an interpretation for which  $A$  is satisfied by at least one sequence in  $\Sigma$ .
- $A$  is contradictory (unsatisfiable) iff  $\neg A$  is logically valid (i.e., iff  $A$  is false for every interpretation).

- $A$  logically implies  $B$  iff, in every interpretation, any sequence satisfying  $A$  also satisfies  $B$ .  
Logical consequence of a set of sequences . . .
- Any sentence of a formal language that is an instance of a logically valid wff is called logically true, and an instance of a contradictory wff is said to be logically false.

## Fundamentals of Quantificational Logic: Deduction

### Proof theory

Axiomatisation in Hilbert style:  
logical ( $\mathcal{P}\mathcal{L}$ ) and nonlogical (theory specific) axioms

*Axioms of  $\mathcal{P}\mathcal{L}$*

$\mathcal{P}\mathcal{L}1 \dots \mathcal{P}\mathcal{L}3 = \mathcal{L}1 \dots \mathcal{L}3$

$\mathcal{P}\mathcal{L}4 \quad \bigwedge_x A(x) \rightarrow A(t); \quad \text{term } t \text{ free for } x \text{ in } A(x)$

$\mathcal{P}\mathcal{L}5 \quad \bigwedge_x (A \rightarrow B) \rightarrow (A \rightarrow \bigwedge_x B);$

$A$  is a wff in which  $x$  does not occur free.

*Inference rules:*

$MP \quad A, A \rightarrow B \vdash B$  (Modus Ponens)

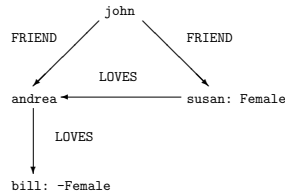
$Gen \quad A \vdash \bigwedge_x A$

## Fundamentals of Quantificational Logic: Metatheory

- Consistency
- Deduction Theorem
- Soundness
- Completeness
- $\vdash S \leftrightarrow \models S$  (proof theory vs. model theory)
- UNDECIDABILITY (semi-decidability)

## Theories and Models in FOL: An Example

$\Gamma = \text{FRIEND}(\text{john}, \text{susan}) \wedge \text{FRIEND}(\text{john}, \text{andrea}) \wedge$   
 $\text{LOVES}(\text{susan}, \text{andrea}) \wedge \text{LOVES}(\text{andrea}, \text{bill}) \wedge \text{Female}(\text{susan}) \wedge$   
 $\neg \text{Female}(\text{bill})$



## Example (cont.)

Does John have a female friend loving a male (i.e. not female) person?

$\Gamma \models \bigvee_{X,Y}. \text{FRIEND}(\text{john}, X) \wedge \text{Female}(X) \wedge \text{LOVES}(X, Y)$   
 $\wedge \neg \text{Female}(Y).$

Answer: **YES**

$\Gamma \not\models \text{Female}(\text{andrea})$   
 $\Gamma \not\models \neg \text{Female}(\text{andrea})$

Unique minimal model:

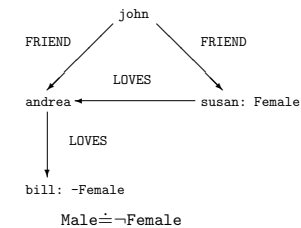
$\Delta^{\mathcal{I}} = \{\text{john}, \text{andrea}, \text{bill}\}$   
 $\text{Female}^{\mathcal{I}} = \{\text{susan}\}$

Two paths:

1.  $\text{FRIEND}(\text{john}, \text{susan}), \text{Female}(\text{susan}),$   
 $\text{LOVES}(\text{susan}, \text{andrea}), \neg \text{Female}(\text{andrea})$
2.  $\text{FRIEND}(\text{john}, \text{andrea}), \text{Female}(\text{andrea}),$   
 $\text{LOVES}(\text{andrea}, \text{bill}), \neg \text{Female}(\text{bill})$

## Example (cont.)

$\Gamma_1 =$   
 $\text{FRIEND}(\text{john}, \text{susan}) \wedge \text{FRIEND}(\text{john}, \text{andrea}) \wedge$   
 $\text{LOVES}(\text{susan}, \text{andrea}) \wedge \text{LOVES}(\text{andrea}, \text{bill}) \wedge$   
 $\text{Female}(\text{susan}) \wedge$   
 $\text{Male}(\text{bill}) \wedge \bigwedge_X. \text{Male}(X) \leftrightarrow \neg \text{Female}(X)$



## Example (cont.)

Does John have a female friend loving a male person?

$$\Gamma_1 \models \bigvee_{X,Y}. \text{FRIEND}(\text{john}, X) \wedge \text{Female}(X) \wedge \text{LOVES}(X, Y) \\ \wedge \text{Male}(Y).$$

$$\Gamma_1 \not\models \text{Female}(\text{andrea}) \quad \Gamma_1 \not\models \neg \text{Female}(\text{andrea})$$

$$\Gamma_1 \not\models \text{Male}(\text{andrea})$$

$$\Gamma_1 \not\models \neg \text{Male}(\text{andrea})$$

Four minimal models, a unique one does not exist:

$$\Delta_1^{\mathcal{I}} = \{\text{john}, \text{susan}, \text{andrea}, \text{bill}\}$$

$$\text{Female}_1^{\mathcal{I}} = \{\text{susan}, \text{andrea}\}$$

$$\text{Male}_1^{\mathcal{I}} = \{\text{bill}, \text{john}\}$$

$$\Delta_1^{\mathcal{I}} = \{\text{john}, \text{susan}, \text{andrea}, \text{bill}\}$$

$$\text{Female}_1^{\mathcal{I}} = \{\text{susan}, \text{andrea}, \text{john}\}$$

$$\text{Male}_1^{\mathcal{I}} = \{\text{bill}\}$$

$$\Delta_2^{\mathcal{I}} = \{\text{john}, \text{susan}, \text{andrea}, \text{bill}\}$$

$$\text{Female}_2^{\mathcal{I}} = \{\text{susan}\}$$

$$\text{Male}_2^{\mathcal{I}} = \{\text{bill}, \text{andrea}, \text{john}\}$$

$$\Delta_2^{\mathcal{I}} = \{\text{john}, \text{susan}, \text{andrea}, \text{bill}\}$$

$$\text{Female}_2^{\mathcal{I}} = \{\text{susan}, \text{john}\}$$

$$\text{Male}_2^{\mathcal{I}} = \{\text{bill}, \text{andrea}\}$$

## Model Checking

Verify that a given interpretation  $\mathcal{I}$  is a model for a closed formula  $\phi$ :

$$\models_{\mathcal{I}} \phi$$

An interpretation is also called a *relational structure*

**Example:**

$$\Delta = \{a, b\}$$

$$P(a)$$

$$Q(b)$$

is a model of the formula

$$\bigvee_y (P(y) \wedge \neg Q(y)) \wedge \bigwedge_z (P(z) \vee Q(z))$$

## Further Reasoning Problems: Subsumption and Instance Checking

### Subsumption

- $\phi \sqsupseteq \psi$  “ $\phi$  subsumes  $\psi$ ”,  
 $\phi$  and  $\psi$  are predicate symbols of same arity
- $\models \bigwedge_{\hat{x}} (\psi(\hat{x}) \rightarrow \phi(\hat{x}))$
- The subsumption relation is a *partial ordering* relation (transitive, reflexive, antisymmetric) in the space of predicates of same arity.

### Instance checking

- The constant  $a$  is an instance of the unary predicate  $P$
- $\Gamma \models P(a)$

## Decidability

Given a logic  $\mathcal{L}$ , a reasoning problem is said to be **decidable**, if there exists a computational process (an algorithm) that solves the problem in a finite number of steps, i.e., the process terminates always.

- The problem of deciding whether a formula  $\phi$  is logically implied by a theory  $\Gamma$  is undecidable in full FOL
- Logical implication is decidable if we restrict it to propositional logic
- Logical implication is decidable if we restrict FOL to use only at most two variable names (=  $\mathcal{L}_2$ )

The problem of (un)decidability is a general property of the problem and not of a particular algorithm solving it.

## Expressive Power

- Some logics can be made decidable by sacrificing some *expressive power*.
- A logical language  $\mathcal{L}_a$  has more expressive power than a logical language  $\mathcal{L}_b$ , if each formula of  $\mathcal{L}_b$  denotes the “same” set of models as its corresponding formula in  $\mathcal{L}_a$ , and if there is a formula of  $\mathcal{L}_a$  denoting a set of models which is denoted by no formula in  $\mathcal{L}_b$ .

*Example:*

Let  $\mathcal{L}_a$  be FOL and  $\mathcal{L}_b$  be FOL without negation and adjunction. Given a common domain, the  $\mathcal{L}_a$  formula  $\forall x (P(x) \vee Q(x))$  has a set of models which cannot be captured by any formula of  $\mathcal{L}_b$ .

## Extensions to Standard FOL

Why do we consider extensions to the standard logical language(s)?

⇒ Requirements of knowledge representation / domain modelling and processing

- Representation of structured domains / structured objects  
Ontological parsimony of standard logic; domains are “flat”  
⇒ Need for *expressive adequateness* and *notational efficiency*
- Object sorts: Multisorted logics

In later chapters:

- Intensional expressions: Create contexts which violate standard principles of logic, e.g. substitution of identities

- Modalities:  
Necessity and possibility  
Epistemic operators (knowledge and belief)
- Temporal logic
- Intensional logic with types for natural language semantics: Montague
- Non-monotonic reasoning and reason maintenance

## Multisorted Logics — A Simple Extension

- First order logics in which more than one sort of individual variables exists;
- are interpreted not over a unique domain, but on a variety of domains;
- can be reduced to one-sorted by suitable one-place predicates and relative quantifiers.

Notation:

$$\bigwedge_{x \in \mathbf{N}} x > 0 \quad \text{instead of} \quad \bigwedge_x \mathbf{N}(x) \rightarrow x > 0$$

Technical reasons for introducing sorts in automatic theorem proving: In a proof, a variable in a wff can be substituted by any term — only unifiability is decisive. Problem: This may cause a *huge* combinatorial overhead.

Solution: Each term in a formula is assigned a class (i.e., a *sort*) and unification is only allowed if both terms are of the same sort.

In programming languages (“types”) and KR, sorts are usually related to each other; in the special case of hierarchical taxonomies we have *partial orders*: “**order-sorted logic**”.

## Relevance Logic

A subjunction  $A \rightarrow B$  shall only exist if antecedent  $A$  is relevant for consequent  $B$ , i.e. if  $B$  is justified relative to the assumption  $A$  ( $A$  is *required*).

Paradoxes of subjunction are caused by the truth functional definition of subjunction:  $A \rightarrow B \iff \neg A \vee B$

Ex-falso-quodlibet of classical and intuitionistic logic trivializes inconsistent theories because it admits the deduction of arbitrary sentences, e.g. Lewis' paradoxes:  $A \rightarrow (\neg A \rightarrow B)$  or  $A, \neg A \prec B$ , resp.

Ex-falso-quodlibet-negatio:  $A \rightarrow (\neg A \rightarrow \neg B)$  or  $A, \neg A \prec \neg B$  (from false follows the falseness of everything), and further paradoxes.