M305

Institute of Distance and Open Learning
Gauhati University

M.A./M.Sc. in Mathematics
Semester 3

Paper V
Mathematical Logic

Contents:

Contributors :

Prof. B. P. Chetiya (Retd) Dept. of Mathematics
Gauhati University

Mr. Priyanka Pratim Baruah Asst. Professor
Dept. of Mathematics, GIMT, Guwahati

Editorial Team :

Prof. Kuntala Patra Dept. of Mathematics
Gauhati University

Prof. Pranab Jyoti Das Director, i/c
IDOL, Gauhati University

Dipankar Saikia Editor, (SLM) GU, IDOL

Cover Page Desi ning:

Mr. Bhaskarjyoti Goswami: IDOL, Gauhati University

 Published on behalf of the Institue of Distance and Open Learning, Gauhati University by Prof. Pranab Jyoti Das, Director, i/c, and printed at under the aegis of G.U. Press as per procedure laid down for the purpose. Copies printed 500.

UNIT : 1

INFORMAL STATEMENT CALCULUS

Introduction :

Logic or at least logical mathematics, consists of deduction. We shall examine the rules of deduction making use of the precision which characterises a mathematical approach. In doing this, if we are to have any precision at all we must make our language unambiguous, and the standard mathematical way of doing that is to introduce a symbolic language, with the symbols having precisely stated meanings and uses.

1.1. Sentential Connectives, Truth-tables

A **statement** or **verbal assertion** or a **truth-functional** is a meaningful sentence having a truth-value i.e. we can say that the assertion made in the sentence is either true or false. **Truth** or **falsity** of a statement is denoted by T or F respectively.

Examples 1.1.1

1. Two plus two is five (F)
2. Delhi is the Capital of India (T)
3. Paris is in Europe (T)
4. Five is greater than six (F)
5. The sun rises in the west (F)

The following are not statements in the sense of logic:

1. Where are you going?
2. I thank you for your kindness
3. Are you coming to the party?

Statements are usually denoted by straight capital letters A,B,C,.... However, we sometimes use small letters like p,q,r,.... to denote statements.

The **negation** of a statement A is a statement having truth-value opposite to that of A. The negation of a statement A is denoted by ~A or $\wedge$ A.

Examples 1.12

1. Let A be the statement "Two plus two is five", i.e. "$2 + 2 = 5$". (Note that "$2 + 2 = 5$" is a statement.)

   Then A is the statement "Two plus two is not five", i.e. "$2 + 2 \neq 5$". The truth value of A is F whereas the truth-value of ~A is T.

2. Let B denote the statement "Delhi is the Capital of India". Then B is the statement "Delhi is not the Capital of India".

   The truth-value of B is T whereas the truth-value of ~B is F.

3. Let C denote the statement "The Prime Minister is honest". Then ~C is the statement "The Prime Minister is dishonest".

4. Let D denote the statement "He is rich". Then ~D is the statement "He is poor".

Any statement A can have a truth-value T or F. The corresponding truth-table of ~A is shown in the following table:

A	~A
T	F
F	T

The table given above is an example of a **truth-table** which will be discussed in more details after a short while from now.

It is to be noted that **negation** of a statement may be regarded as an operation on statments. This operation can be performed on a **single** statement, and so we call it a **unary operation** on the set of statements. There are four other operations by which we can combine two statements to get a new statement. A new statement obtained in this way is called a **composite statement** or a **compound statement**. The negation of a statement is also regarded as a composite or compound statement.

Examples 1.1.3 (Of composite statements):

1. Delhi s in India **and** Lahore is in Pakistan.
2. Paris is in France **and** London is in Canada.
3. 2+2=4 **and** 3+3=5.
4. Calcutta is in India **or** 2+4=7.
5. Guwahati is in Assam **or** Calcutta is in West Bengal.
6. London is in India **or** 3+3=7.
7. If the function f is derivable at x, **then** f is continuous at x.
8. If 2+2=5, **then** 3+7=10.
9. If 3+3=7, **then** 3+8=12.
10. Paris is in France **if and only if** 2+2=5.
11. Paris is in France **if and only if** 2+2=4.
12. Paris is in England **if and only if** 2+2=5.

Now our aim is to discuss how to determine the truth-value (T or F) of a composite statement. For this we need some definitions and symbols relating to composite statements.

In contrast to a composite statement, a statement of the sim plest type, presented earlier, will be called **Atomic Statement**. Letters like A,B,C,...., p,q,r,... are used to denote atomic statements only, and such letters are called **statement letters.**

Conjunction (∧) and Disjunction(∨):

The operation by which we combine two atomic statements by using **'and'** is called **conjunction,** as in examples 1,2,3. If we denote the first statement by A and the second statement by B then the composite statement "A and B" is denoted by $A \wedge B$.

The operation by which we combing two atomic statements by using **'or' is called disjunction,** as in Examples 4, 5 and 6. For two statements A and B the composite statement "A or B" is denoted by $A \vee B$. Note that, logically "A or B" means "A or B or both A and B". This meaning of A or B is called the **inclusive meaning.** However in ordinary usage of 'or', A or B means only one of A and B but **not both.** This meaning of A or B is called the **exclusive meaning.** In logic A or B will always have the **inclusive meaning.**

Conditional (→) and biconditional (↔) operations:

Let A and B be two atomic statements. The operation on A and B which gives a composite statement of the form "if A, then B" is called a **conditional operation.** Such a composite statement is called a **conditional statement** and is denoted by $A \rightarrow B$. The statements of examples 7,8,9 are of this type. The operation on A and B which gives a composite statement of the form "A if and only if B" or in short "A iff B" is called a **biconditional operation.** Such a composite statement is called a **biconditional statement** and is denoted by $A \leftrightarrow B$. Note that "A if and only if B" is equivalent to "If A, then B **and** if B then A" which can now be expressed in symbols as

$$(A \rightarrow B) \wedge (B \rightarrow A)$$

The symbols $\sim, \wedge, \vee, \rightarrow, \leftrightarrow$ are called **sentential connectives or propositional connectives** or simply **connectives** of logic. The first symbol ~ is a **unary connective** in the sense that it involves a **single** atomic statement. The other four symbols, namely $\wedge, \vee, \rightarrow$, and $\leftrightarrow$ are **binary connectives** in the sense that each one of them involves a **pair** of atomic statements.

Any composite statement which contains atomic statements in it can be symbolised by using statement letters and the five connectives mentioned above.

Examples 1.1.4

Consider the following composite statements:

1. Milk is black and water is liquid.
2. Milk is not black and water is liquid.
3. Milk is black or water is liquid.
4. If water is liquid, then milk is black.
5. Milk is black if and only if water is not liquid.

Only two atomic statements occur in the above examples. These are:

Milk is black.

Water is liquid.

Let us use statement letters A and B respectively for these two atomic statements:

A: Milk is black.

B. Water is liquid.

Then the five statements mentioned above can be symbolised as:

1. $A \wedge B$
2. $(\sim A) \wedge B$
3. $A \vee B$
4. $B \rightarrow A$
5. $A \leftrightarrow B$

The symbolic forms of the five **fundamental types (simplest types)** of composite statements are as follows:

(i)	$\sim A$	negation
(ii)	$A \wedge B$	conjunction
(iii)	$A \vee B$	disjunction
(iv)	$A \rightarrow B$	conditional
(v)	$A \leftrightarrow B$	biconditional

where A and B denote atomic statements.

Sometimes, for the sake of simplicity in the use of language, we do not distinguish between an **atomic statement** and the **statement letter denoting the atomic statement**. For example, we may say "Let A be the statement 'Delhi is in India' " instead of saying "Let A be the statement letter denoting the statement 'Delhi is in India' ". Also when there is no scope for any ambiguity we may write simply 'statement' for 'atomic statement' as we have done just above.

We now turn our attention to determination of truth-values of fundamental composite statements which can be symbolised as $\sim A$, $A \wedge B$, $A \vee B$, $A \rightarrow B$, $A \leftrightarrow B$ where A and B are atomic statements. Note that as in the case of atomic statements, we do not make any distinction between a composite statement and the symbolic expression for it. The same convention will be followed for **more complicated composite statements** which will have to be dealt with in future.

In $\sim A$, A can have two truth-value assignments, T and F. In each of $A \wedge B$, $A \vee B$, $A \rightarrow B$, $A \leftrightarrow B$, A and B together can have **four** truth-value assignments, namely (T,T), (F,T), (T,F) and (F,F). Keeping these things in mind five tables, called **truth-tables,** are constructed which determine the truth-values of the fundamental composite statements. In each table an entry in the column on the extreme right gives the truth-value of the composite statement corresponding to the truth-value assignment to the atomic statement(s) shown by the other entry (entries) in the same row.

1. **Truth-table for $\sim A$**

A	$\sim A$
T	F
F	T

2. **Truth-table for $A \wedge B$**

A	B	$A \wedge B$
T	T	T
F	T	F
T	F	F
F	F	F

A B is true if both the components of A and B are true, and false if any one of the components is false.

3. Truth-table for $A \vee B$

A	B	$A \vee B$
T	T	T
F	T	T
T	F	T
F	F	F

A B is true if any one of the components A and B is true, and false only when both the components A and B are false.

4. True-table for $A \rightarrow B$

A	B	$A \rightarrow B$
T	T	T
F	T	T
T	F	F
F	F	T

A B is false only when A is true and B is False, and in all other case A B is true

5. True-table for $A \leftrightarrow B$

A	B	$A \leftrightarrow B$
T	T	T
F	T	F
T	F	F
F	F	T

$A \leftrightarrow B$ is true only when both A and B are true or both A and B are false. $A \leftrightarrow B$ is false if one of A and B is true and the other false.

The above tables are usually referred to as the truth-tables for the connectives $\sim, \wedge, \vee, \rightarrow, \leftrightarrow$. We now demonstrate with examples how the above tables can be used to determine truth-values of composite statements.

Examples 1.1.5

Consider the following composite statements

1. $2+2=4$ and $2\times3=6$
2. $2+2=5$ and $2\times3=6$
3. $2+2=5$ and $2\times3=9$
4. Milk is black and water is liquid
5. $2+2=4$ and $2\times3=6$
6. $2+2=5$ and $2\times3=9$
7. $2+2=5$ and $2\times3=6$

8. Milk is black or water is liquid
9. If 1+1=2, then Paris is in France.
10. If $2+2 \neq 4$, then 3+3=9
11. If milk is black, then water is liquid
12. 2+2=4 if and only if $2\times3=6$
13. $2+2\neq4$ if and only if $2\times3=6$
14. $2+2\neq4$ if and only if $2\times3=6$
15. Milk is black if and only if water is solid.

Each of the above composite statements contains two atomic statements. Let us denote them by A and B in the order of occurrence, i.e. A for the **first statement** and B for the **second.** The given composite statements can now be symbolised in terms of A, B and the connectives. Each composite statement is of a form $A \wedge B$ or $A \vee B$ or $A \rightarrow B$ or $A \leftrightarrow B$. Moreover, in each case the truth-values of A and B are **known.** Therefore the truth-values of the given composite statements can be determined by using the tables 2,3,4,5. We now demonstrate the truth-values of these statements by (partial) truth-tables:

1.

A	B	$A \wedge B$
T	T	T

The statement is **true.**

2.

A	B	$A \vee B$
F	T	F

The statement is **false.**

3.

A	B	$A \wedge B$
F	F	F

The statement is false.

4.

A	B	$A \wedge B$
F	T	F

The statement is **false.**

5.

A	B	$A \vee B$
T	T	T

The statement is **true.**

6.

A	B	$A \vee B$
F	F	F

The statement is **false.**

7.

A	B	$A \vee B$
F	T	T

The statement is **true.**

8.

A	B	$A \vee B$
F	T	T

The statement is **true.**

9.	A	B	$A \to B$
	T	T	T

The statement is **true.**

10.	A	B	$A \to B$
	F	F	T

The statement is **true.**

11.	A	B	$A \to B$
	F	T	T

The statement is **true.**

12.	A	B	$A \leftrightarrow B$
	T	T	T

The statement is **true.**

13.	A	B	$A \leftrightarrow B$
	F	T	F

The statement is **false.**

14.	A	B	$A \leftrightarrow B$
	F	F	T

The statement is **true.**

15.	A	B	$A \leftrightarrow B$
	F	F	T

The statement is **true.**

Very often we come across much more complex composite statements involving several atomic statements and several connectives. Such statements can also be symbolised by using the symbolism we have developed so far.

Example 1.1.6 Let A be 'He is tall' and B be 'He is handsome'. Let us symbolise the following statements:

(a) He is tall but not handsome.

(b) It is false that he is short or handsome.

(c) He is neither tall nor handsome.

(d) He is tall or he is short and handsome.

(e) It is not true that he is short or not handsome.

Sometimes it becomes easier to symbolise a given statement if it is transformed into a convenient form.

(a) The statement is equivalent to:

He is tall and he is not handsome.

So it can be symbolised as

$A \wedge (\sim B)$

(b) The statement is the **negation** of the statement:

He is short or he is handsome.

i.e. **He is not tall or he is handsome.**

So the given statement can be symbolised as

$\sim((\sim A) \vee B)$

(c) The statement is equivalent to:

He is not tall and he is not handsome

So the given statement canbe symbolised as

$(\sim A) \wedge (\sim B)$

(d) The statement is disjunction of the following two statements:

He is tall

He is not tall and he is handsome.

So, we can symbolise the given statement as $A\wedge((\sim A)\wedge B)$

(e) The statement is the negation of the disjunction of the following two statements:

He is not tall

He is not handsome

So a symbolic expression of the given statement is

$\sim A((\sim A)\vee(\sim B)$

Example 1.1.7 Let p be 'He is rich' and q be 'He is happy'. Consider the following statements:

(1) He is neither rich nor happy.
(2) He is poor but happy.
(3) He cannot be both rich and happy.
(4) If he is unhappy, then he is poor.
(5) If he is not poor and happy, then he is rich.
(6) To be rich means the same as to be happy.
(7) He is poor or else he is both rich and unhappy.

The above statements are in everyday language. As we saw in the last example, it would be convenient to symbolise them if we would first express them in rather 'mathematical' forms. For each statement we give below an alternative form with the corresponding symbolic expression.

(1) He is not rich and he is not happy.

$(\sim p)\wedge(\sim q)$

(2) He is not rich and he is happy.

$(\sim p)\wedge q$

(3) It is not true that he is rich and he is happy.

$\sim(p\wedge q)$

(4) If he is not happy, then he is not rich.

$(\sim p)\rightarrow(\sim q)$.

(5) If he is not poor and he is happy, then he is rich.

$((\sim p)\wedge q)\rightarrow p$.

(6) He is rich if and only if he is happy.

$p\leftrightarrow q$

(7) He is not rich, or he is rich and he is not happy.

$(\sim p)\vee(p\wedge(\sim q))$.

A symbolic expression of a statement is usually called a **statement form**. Speaking more formally, an expression built up from a number of statement letters A, B, C... by appropriate application of the connectives $\sim, \wedge, \vee, \rightarrow, \leftrightarrow$ is called a **statement form.** A statement form is also called a **proposition** or a **boolean polynomial.** The name **Propositional Calculus** originates from the word **proposition** used in this sense-Propositional Calculus is the analysis of the truth-values of propositions. Note that Propositional Calculus is called **Statement Calculus** also.

We usually denote statement forms by curly capital letters $\mathcal{A},\mathcal{B},\mathcal{C}$... However, statement letters A,B,C,... are also regarded as statement forms.

Note that a statement letters are given a statement form may be translated into everyday language.

Example 1.1.8 Let p be 'H eis rich' and let q be 'He is happy'. Consider the following statement forms:

(1) $q\leftrightarrow(\sim p)$ (2) $p\vee(\sim q)$

(3) $(\sim p)\rightarrow q$ (4) $((\sim p)\wedge q)$

These statement forms may be translated into language as follows:-

(1) He is happy if and only if he is not rich.

or, in everyday language

To be happy means the same as to be poor.

(2) He is rich or he is unhappy.

(3) If he is not rich, then he is happy.

or, in everyday language

To be poor is to be happy.

(4) If he is not rich but happy, then he is rich.

or, in everyday language

To be poor but happy means to be rich.

Truth-tables for Statement Forms

The truth-value of a statement form in the statement letters A,B,C,... can be determined from the truth-values of A,B,C,...by constructing appropriate truth-tables. In the construction of such a truth-table we make use of the truth-tables for the fundamental composite statements, i.e. the truth-tables for the connectives $\sim, \wedge, \vee, \rightarrow, \leftrightarrow$ namely.

A	~A
T	F
F	T

A	B	A∧B
T	T	T
F	T	F
T	F	F
F	F	F

∧

A	B	A∨B
T	T	T
F	T	T
T	F	T
F	F	F

∨

A	B	A→B
T	T	T
F	T	T
T	F	F
F	F	T

→

A	B	A↔B
T	T	T
F	T	F
T	F	F
F	F	T

↔

Note that in the above tables we can replace the statement letters A,B by any statement forms $\mathcal{A}$ and $\mathcal{B}$.

In constructing a truth-table for a statement form we **plot** the truth-values 'component by component' and 'step by step' as domenstrated in the following examples.

Example 1.1.9 The truth-table for the statement form ~(A∧(~B)) is as follows:

A	B	~B	A∧(~B)	~(A∧(~B))
T	T	F	F	T
F	T	F	F	T
T	F	T	T	F
F	F	T	F	T

Example 1.1.10 The truth-table for the statement form ((A)B)C is as follows:

A	B	C	~A	(~A)∨(B)	((~A)∨B)→C
T	T	T	F	T	T
T	F	T	F	F	T
T	T	F	F	T	F
T	F	F	F	F	T
F	T	T	T	T	T
F	F	T	T	T	T
F	T	F	T	T	F
F	F	F	T	T	F

Such truth-tables for statement forms are very useful in the sense that these give the truth-values of the statement forms corresponding to the truth-values of the statement letters. In other words, if we know the truth-values of the statement letters we immediately know the truth-value of the statement form, i.e. the truth-value of the statement represented by the statement form.

Concise way of construction of truth-tables

Truth-tables of statement forms are usually constructed in a concise way. For example, the truth-tables in the above examples, i.e., the truth-tables for the statement forms ~(A∧(~B)) and ((~A)∨B)→C may be constructed as follows:

~	(A	∧	(~	B))
T	T	F	F	T
T	F	F	F	T
F	T	T	T	F
T	F	F	T	F
4	1	3	2	1

((~	A)	∨	B)	→	C
F	T	T	T	T	T
F	T	F	F	T	T
F	T	T	T	F	F
F	T	F	F	T	F
T	F	T	T	T	T
T	F	T	F	T	T
T	F	T	T	F	F
T	F	T	F	F	F
2	1	3	1	4	1

The numbers at the bottom of the columns of the tables idicate the order in which the columns are to be filled up. Columns marked 1 are to be filled up first, and go on filling up the columns marked 2,3,4 and so on. The entries in the final column (column 4 in the above examples) give the truth-values of the statement form under consideration, i.e. an entry in the final column is the truth-value of the statment form corresponding to the truth-values of the statement letters occurring in the same row in which the entry occurs.

We already mentioned that a statement form may be regarded as a function f(A,B,C,...) in the statement letters A,B,C,.... Now we see that such a function may take the truth-value T or F. In this sense a statement form is sometimes referred to as a **truth-function.**

1.2. Tautology, Contradiction, Contingent

A statement form which is **always true**, no matter what the truth-values of its statement letters may be, is called a **tautology**. In other words, a statement form is a tautology if and only if its corresponding truth-function takes the value T only, or equivalently, if and only if the final column in its truth-table contains only Ts.

Example 1.2.1 The following statement forms are tautologies:

(i) A∨(~A) (ii) ~(A∧(~A))

(iii) (A∧B)→A (iv) A→(A∨B)

(v) ((p→q)∧(q→r))→(p→r).

When we construct the truth-tables we see that the final columns contain only Ts.

(i)

A	∨	(~	A)
T	T	F	T
F	T	T	F
1	3	2	1

This tautology is referred to as the **Law of Excluded Middle.**

(ii)

~	(A	∧	(~	A))
T	T	F	F	T
T	F	F	T	F
4	1	3	2	1

(iii)

(A	∧	B)	→	A
T	T	T	T	T
F	F	T	T	F
T	F	F	T	T
F	F	F	T	F
1	2	1	3	1

(iv)

A	→	(A	∨	B)
T	T	T	T	T
F	T	F	T	T
T	T	T	T	F
F	T	F	F	F
1	3	1	2	1

(v)

((p	→	q)	∧	(q	→	r))	→	(p	→	r)
T	T	T	T	T	T	T	T	T	T	T
T	F	F	F	F	T	T	T	T	T	T
T	T	T	F	T	F	F	T	T	F	F
T	F	F	F	F	T	F	T	T	F	F
F	T	T	T	T	T	T	T	F	T	T
F	T	F	T	F	T	T	T	F	T	T
F	T	T	F	T	F	F	T	F	T	F
F	T	F	T	F	T	F	T	F	T	F
1	2	1	3	1	2	1	4	1	2	1

A statement form which is always **false**, no matter what the truth-values of its statement letters may be is called a **contradiction.** In other words a statement form is a contradiction if and only if its corresponding truth-function takes the value F only, or equivalently, if and only if the final column in its truth-table contains only Fs.

Note that the negation of a contradiction is a tautology and vice-versa.

Example 1.2.2 The following statement forms are contradictions

(i) $p\wedge(\sim p)$ (ii) $(p\wedge q)\wedge(\sim(p\vee q))$

(i)

p	∧	(~	p)
T	F	F	T
F	F	T	F
1	3	2	1

(ii)

(p	∧	q)	∧	(~	(p	∨	q))
T	T	T	F	F	T	T	T
F	F	T	F	F	F	T	T
T	F	F	F	F	T	T	F
F	F	F	F	T	F	F	F
1	2	1	4	3	1	2	1

If the final column of the truth-table of a statement form contains both Ts and Fs, then the statement form is called a **contingent.** Thus, a contingent is a statement form whose truth-value **depends** on the individual truth-value assignments to the statement letters.

Example 1.2.3 (~A∧(~B)) is a contingent.

~	(A	∧	(~	B))
T	T	F	F	T
T	F	F	F	T
F	T	T	T	F
T	F	F	T	F
4	1	3	2	1

NOTATION If $\mathcal{A}$ is a tautology, we write $\mathcal{A}$.

Definitions : If $\mathcal{A}\rightarrow\mathcal{B}$ is a tautology, then we say that $\mathcal{A}$ **logically implies** $\mathcal{B}$ or that $\mathcal{B}$ **is a logical consequence** of $\mathcal{A}$.

If $\mathcal{A}\leftrightarrow\mathcal{B}$ is a tautology then $\mathcal{A}$ and $\mathcal{B}$ are said to be **logically equivalent**, and written $\mathcal{A}\equiv\mathcal{B}$.

Example 1.2.4 A∧B logically implies A

We have to check that (A∧B)→ A is a tautology:

(A	∧	B)	→	A
T	T	T	T	T
F	F	T	T	F
T	F	F	T	T
F	F	F	T	F
1	2	1	3	1

Example 1.2.5 B is a logical consequence of A∧(A→B)

We have to check that (A∧(A→B))→B is a tautology.

(A	∧	(A	→	B))	→	B
T	T	T	T	T	T	T
F	F	F	T	T	T	T
T	F	T	F	F	T	F
F	F	F	T	F	T	F
1	3	1	2	1	4	1

Example 1.2.6 $A\leftrightarrow B$ is logically equivalent to $(A\rightarrow B)\wedge(B\rightarrow A)$.

We check that $(A\leftrightarrow B)\leftrightarrow((A\rightarrow B)\wedge(B\rightarrow A))$ is a tautology

(A	↔	B)	↔	((A	→	B)	∧	(B	→	A))
T	T	T	T	T	T	T	T	T	T	T
F	F	T	T	F	T	T	F	T	F	F
T	F	F	T	T	F	F	F	F	T	T
F	T	F	T	F	T	F	T	F	T	F
1	2	1	4	1	2	1	3	1	2	1

Example 1.2.7 $(\sim A)\vee B$ and $(\sim B)\vee A$ are **not** logically equivalent.

We check that $((\sim A)\vee B)\leftrightarrow((\sim B)\vee A)$ is **not** a tautology:

((~	A)	∨	B)	↔	((~	B)	∨	A)
F	T	T	T	T	F	T	T	T
T	F	T	T	F	F	T	F	F
F	T	F	F	F	T	F	T	T
T	F	T	F	T	T	F	T	F
2	1	3	1	4	2	1	3	1

In the symbolic expressions for statement forms we have so far made extensive use of parentheses (small brackets) with the obvious implication that the operations inside the parentheses are to be performed first. In order to **minimise** the use of parentheses we follow a convention. However, we cannot totally avoid use of parentheses. The convention we are going to follow hence forward is:

First, dispose of the parentheses part(s), if any.

Secondly, perform the operations of the connectives in the order

$$\sim,\ \wedge,\ \vee,\ \rightarrow,\ \leftrightarrow.$$

Now using this convention the statement forms

$$A\vee(\sim A)$$

$$\sim(A\wedge(\sim A))$$

$$A\rightarrow(A\vee B)$$

$$((p\rightarrow q)\wedge(q\rightarrow r))\rightarrow(p\rightarrow r)$$

may be expressed as

$$A\vee\sim A$$

$$\sim(A\wedge\sim A)$$

$$A\rightarrow A\vee B$$

$$(p\rightarrow q)\wedge(q\rightarrow r)\rightarrow(p\rightarrow r)$$

respectively.

Laws of Logical Equivalence

A short while ago we wrote about logical equivalence of statement forms. We give below a set of **laws of logical equivalence** for any three statement forms $\mathcal{A}, \mathcal{B}, \mathcal{C}$, any tautology $\mathcal{T}$ and any contradiction $\mathcal{T}$. Recall the symbol $\equiv$ for equivalence of two statement forms.

Idempotent Laws

1(a) $\mathcal{A} \vee \mathcal{A} \equiv \mathcal{A}$ 1(b) $\mathcal{A} \wedge \mathcal{A} \equiv \mathcal{A}$

Associativity Laws

2(a) $(\mathcal{A} \vee \mathcal{B}) \vee \mathcal{C} \equiv \mathcal{A} \vee (\mathcal{B} \vee \mathcal{C})$ 2(b) $(\mathcal{A} \wedge \mathcal{B}) \wedge \mathcal{C} \equiv \mathcal{A} \wedge (\mathcal{B} \wedge \mathcal{C})$

Commutativity Laws

3(a) $\mathcal{A} \vee \mathcal{B} \equiv \mathcal{B} \vee \mathcal{A}$ 3(b) $\mathcal{A} \wedge \mathcal{B} \equiv \mathcal{B} \wedge \mathcal{A}$

Distributivity Laws

4(a) $\mathcal{A} \vee (\mathcal{B} \wedge \mathcal{C}) \equiv (\mathcal{A} \vee \mathcal{B}) \wedge (\mathcal{A} \vee \mathcal{C})$ 4(b) $\mathcal{A} \wedge (\mathcal{B} \vee \mathcal{C}) \equiv (\mathcal{A} \wedge \mathcal{B})(\mathcal{A} \wedge \mathcal{C})$

Identity Laws

5(a) $\mathcal{A} \vee \mathcal{F} \equiv \mathcal{A}$ 5(b) $\mathcal{A} \wedge \mathcal{T} \equiv \mathcal{A}$

6(a) $\mathcal{A} \vee \mathcal{T} \equiv \mathcal{T}$ 6(b) $\mathcal{A} \wedge \mathcal{F} \equiv \mathcal{F}$

Complement Laws

7(a) $\mathcal{A} \vee \sim\mathcal{A} \equiv \mathcal{T}$ 7(b) $\mathcal{A} \wedge \sim\mathcal{A} \equiv \mathcal{F}$

8(a) $\sim\sim\mathcal{A} \equiv \mathcal{A}$ 8(b) $\sim\mathcal{T} \equiv \mathcal{F}$, $\sim\mathcal{F} \equiv \mathcal{T}$

Demorgan's Laws

9(a) $\sim(\mathcal{A} \vee \mathcal{B}) \equiv \sim\mathcal{A} \wedge \sim\mathcal{B}$ 9(b) $\sim(\mathcal{A} \wedge \mathcal{B}) \equiv \sim\mathcal{A} \vee \sim\mathcal{B}$

We now present some results of **general nature** in the form of **Propositions.** These **Propositions** will give an indication of the oncoming results.

Proposition 1.2.1 If $\mathcal{A}$ and $\mathcal{A} \rightarrow \mathcal{B}$ are tautologies then so is $\mathcal{B}$.

Proof: Since $\mathcal{A}$ is a tautology $\mathcal{A}$ takes the truth-value T only for every truth assignment to the statement letters of $\mathcal{A}$.

$\mathcal{A} \rightarrow \mathcal{B}$ is also a tautology and so $\mathcal{A} \rightarrow \mathcal{B}$ takes the truth-value T only. Therefore, if $\mathcal{B}$ has the truth-value F, then (by the truth-table for $\rightarrow$) $\mathcal{A}$ must have the truth-value F. But this gives a contradiction. Hence $\mathcal{B}$ must have the truth-value T, i.e. $\mathcal{B}$ is a tautology.

Proposition 1.2.2 (Principle of substitution) If $\mathcal{A}$ is a tautology containing the statement letters A_1, A_2,...A_n and $\mathcal{B}$ arises from $\mathcal{A}$ by substituting statement forms $\mathcal{A}_1$, $\mathcal{A}_2$,...,$\mathcal{A}_n$ for A_1, A_2, ..., A_n respectively, then $\mathcal{B}$ is a tautology (i.e. substitution in a tautology yields a tautology).

Before we give the proof of the above proposition we 'illustrate' the result with an eample.

Example 1.2.8 Consider the tautology (Check it!) : $A_1 \wedge A_2 \rightarrow A_1$

Assume $\mathcal{A}$ to be $A_1 \wedge A_2 \rightarrow A_1$. Replace A_1 and A_2 by $\mathcal{A}_1 = B \vee C$ and $\mathcal{A}_2 = C \vee D$ respectively. Then $(B \vee C) \wedge (CD) \rightarrow B \vee C$ is a tautology. This can be checked by constructing a truth-table.

(B	∨	C)	∧	(C	∨	D)	→	B	∨	C
T	T	T	T	T	T	T	T	T	T	T
T	T	F	T	F	T	T	T	T	T	F
T	T	T	T	T	T	F	T	T	T	T
T	T	F	F	F	F	F	T	T	T	F
F	T	T	T	T	T	T	T	F	T	T
F	F	F	F	F	T	T	T	F	F	F
F	T	T	T	T	T	F	T	F	T	T
F	F	F	F	F	F	F	T	F	F	F
1	2	1	3	1	2	1	4	1	2	1

Proof of Proposition 1.1.2

$\mathcal{A}$ is a tautology; and so $\mathcal{A}$ takes the truth-value T only. For any assignment of truth-values to the statement letters in $\mathcal{B}$, suppose the statement forms $\mathcal{A}_1$, $\mathcal{A}_2$,...,$\mathcal{A}_n$ have the truth-values x_1, x_2,...x_n where each x_i is either T or F. If we assign the values x_1, x_2,...x_n to A_1, A_2,...,A_n respectively, the resulting truth-value of $\mathcal{A}$ is same as the truth-value of $\mathcal{B}$ for the given assignment of truth-values. But the truth-value of $\mathcal{A}$ is always T. So the truth-value of $\mathcal{A}$ is also always T, i.e. $\mathcal{B}$ is a tautology.

Proposition 1.2.3 Suppose $\mathcal{A}$ is a statement form which contains another statement form $\mathcal{A}_1$. Suppose $\mathcal{B}$ is obtained from $\mathcal{A}$ by substituting $\mathcal{B}_1$ for one or more occurences of $\mathcal{A}_1$ in $\mathcal{A}$. Then $(\mathcal{A}_1 \leftrightarrow \mathcal{B}_1) \rightarrow (\mathcal{A} \leftrightarrow \mathcal{B})$ is a tautology. Hence if $\mathcal{A}_1$ and $\mathcal{B}_1$ are logically equivalent then so are $\mathcal{A}$ and $\mathcal{B}$.

Proof: Consider assignment of truth-values to the statement letters.

Case 1. Suppose $\mathcal{A}_1$ and $\mathcal{B}_1$ have **opposite** truth-values. Then $\mathcal{A}_1 \leftrightarrow \mathcal{B}_1$ takes the truth-value F.

On the other hand the truth-value of $\mathcal{A} \leftrightarrow \mathcal{B}$ is either T or F. Since $\mathcal{A}_1 \leftrightarrow \mathcal{B}_1$ takes value F, $(\mathcal{A}_1 \leftrightarrow \mathcal{B}_1) \rightarrow (\mathcal{A} \leftrightarrow \mathcal{B})$ takes the value T for both the values T and F of $\mathcal{A} \leftrightarrow \mathcal{B}$. Therefore

$(\mathcal{A}_1 \leftrightarrow \mathcal{B}_1) \rightarrow (\mathcal{A} \leftrightarrow \mathcal{B})$ is a tautology.

Case II. Suppose $\mathcal{A}_1$ and $\mathcal{B}_1$ have the **same** truth-value. Then $\mathcal{A}$ and $\mathcal{B}$ also have the same truth-value, because $\mathcal{B}$ differs from $\mathcal{A}$ only in containing $\mathcal{B}_1$ in some places where $\mathcal{A}$ contains $\mathcal{B}_1$. So the truth-values of both $(\mathcal{A}_1 \leftrightarrow \mathcal{B}_1)$ and $(\mathcal{A} \leftrightarrow \mathcal{B})$ are T, and so the truth-value of.

$(\mathcal{A}_1 \leftrightarrow \mathcal{B}_1) \rightarrow (\mathcal{A} \leftrightarrow \mathcal{B})$ is T

i.e. $(\mathcal{A}_1 \leftrightarrow \mathcal{B}_1) \rightarrow (\mathcal{A} \leftrightarrow \mathcal{B})$ is a tautology.

If $\mathcal{A}$ and $\mathcal{B}$ are logically equivalent, then $\mathcal{A}_1 \leftrightarrow \mathcal{B}_1$ takes the value T. Now since $(\mathcal{A}_1 \leftrightarrow \mathcal{B}_1) \rightarrow (\mathcal{A} \leftrightarrow \mathcal{B})$ is a tautology, $(\mathcal{A} \leftrightarrow \mathcal{B})$ takes the value T, i.e. $\mathcal{A}$ and $\mathcal{B}$ are logically equivalent.

1.3 Adequate Sets of connectives

We mentioned earlier that a statement form in statement letters A,B,C,... may be regarded as a function f(A,B,C,...) where the variables A,B,C,.... can take the values T and F, and the function f also can take the values T and F depending on the values of the variables. In general we can therefore conceive of a function $f(x_1,x_2,...,x_n)$ in n variables $x_1,x_2,....,x_n$ such that each variable x_i as well as the function f can take values from the set {T,F}. Such a function will be referred to as a **truth-function.** Any statement form with n statement letters **generates** (i.e. gives rise to) a truth-function in the sense that if we replace the n statement letters $A_1,A_2,...,A_n$ by the n variables $x_1,x_2,...x_n$ respectively in the statement form then we get a truth-function.

Suppose we are given a truth-function $f(x_1,x_2,...,x_n)$. Can we always construct a statement form $\mathcal{A}$ in the statement letters $A_1,A_2,...,A_n$ which generate the truth-function $f(x_1,x_2,...,x_n)$? The answer is YES. There is a standard technique to do this, and only three connectives, namely ~, ∧, ∨, are sufficient or **adequate** for the purpose. First, we illustrate the technique by some examples.

Example 1.3.1 Let $f(x_1,x_2)$ be a truth-function given by the following 'truth-table':

x_1	x_2	$f(x_1,x_2)$
T	T	F
F	T	T
T	F	T
F	F	T

Table 1

Now our problem is to construct a statement from $\mathcal{D}$ in two statement letters A_1 and A_2 such that $\mathcal{D}$ generates f, i.e. $\mathcal{D}$ has a truth-table identical with the truth-table for f (Table 1), i.e. $\mathcal{D}$ has a truth-table

A_1	A_2	$\mathcal{D}$
T	T	F
F	T	T
T	F	T
F	F	T

Table 2

Such a statement form $\mathcal{D}$ can always be constructed using only **three** connectives ~, ∧, ∨ i.e.~, ∧, ∨ constitute an adequate set of connectives for such a construction.

Table 1 contains $4(=2^2)$ rows. Call them R_1, R_2, R_3, R_4. In each row there are three entries; the extreme right entry gives the truth-value of the function. Pick up those rows in which the extreme right entries are Ts.

These are R_2, R_3, R_4. For each these rows build a conjunction $\mathcal{D}_j$ (j=2,3,4) as follows

$$\mathcal{D}_j = R_j^1 \wedge R_j^2$$

where

$$R_j^i \begin{cases} = A_i \text{ if the entry in the ij-th position of the Table is T} \\ = \sim A_i \text{ if the entry in the ij - th position of the Table is F} \end{cases}$$

i.e. $R_2^1 = A_1,\ R_2^2 = A_2$

$R_3^1 = A_1, R_3^2 = \sim A_2$

$R_4^1 = A_1, R_4^2 = A_2$

and $\mathcal{D}_2 = \sim A_1 \wedge A_2$

$\mathcal{D}_3 = A_1 \wedge \sim A_2$

$\mathcal{D}_4 = \sim A_1 \wedge \sim A_2$

Then construct the disjunction $\mathcal{D}$ of $\mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4$ i.e.

i.e $\mathcal{D} = \mathcal{D}_2 \vee \mathcal{D}_3 \vee \mathcal{D}_4$

i.e $\mathcal{D} = (\sim A_1 \wedge A_2) \vee (A_1 \wedge \sim A_2) \vee (\sim A_1 \wedge A_2)$.

We claim that $\mathcal{D}$ is the statement form that generates $f(x_1,x_2)$. The following observations substantiate our claim.

1. If we give the assignment of R_1 i.e. T,T, to the statement letters A_1, A_2, then all of $\mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4$ are false (F), and therefore $\mathcal{D}$ is false (F).
2. If we give the assignment of R_2 or R_3 or R_4 to A_1, A_2, then $\mathcal{D}_2$ or $\mathcal{D}_3$ or $\mathcal{D}_4$ is true (T) and therefore $\mathcal{D}$ is true (T).

This shows that the truth-table for $\mathcal{D}$ is identical with that of $f(x_1, x_2)$.

Example 1.3.3 To find a statement form in the connectives $\sim$, $\wedge$ and $\vee$ that generates the following truth-function.

x_1	x_2	x_3	$f(x_1,x_2,x_3)$
T	T	T	T
T	F	T	T
T	T	F	T
T	F	F	F
F	T	T	F
F	F	T	F
F	T	F	F
F	F	F	T

We pick up first, second, third and eighth rows in each of which the last entry (i.e. the truth-value of the function) is T. Then we construct $\mathcal{D}_1$, $\mathcal{D}_2$, $\mathcal{D}_3$ and $\mathcal{D}_8$ as follows.

$$\mathcal{D}_1 = A_1 \wedge A_2 \wedge A_3$$
$$\mathcal{D}_2 = A_1 \wedge \sim A_2 \wedge A_3$$
$$\mathcal{D}_3 = A_1 \wedge A_2 \wedge \sim A_3$$
$$\mathcal{D}_8 = A_1 \wedge \sim A_2 \wedge \sim A_3$$

Then let $\mathcal{D} = \mathcal{D}_1 \vee \mathcal{D}_2 \vee \mathcal{D}_3 \vee \mathcal{D}_8$. We claim that $\mathcal{D}$ generates the given truth-function $f(x_1,x_2,x_3)$:

1. If we give the assignment of the fourth or fifth or sixth or seventh row (i.e. the rows in which the last entrie are F) to the statement letters A_1, A_2, A_3, then all of $\mathcal{D}_1$, $\mathcal{D}_2$, $\mathcal{D}_3$, $\mathcal{D}_8$ are false (F), and so $\mathcal{D}$ is also false (F).

2. If we give the assignment of the first or second or third or eighth row (i.e. the rows in which the last entries are T) to the statement letters A_1, A_2, A_3, then $\mathcal{D}_1$ or $\mathcal{D}_2$ or $\mathcal{D}_3$ or $\mathcal{D}_8$ is true (T), and so $\mathcal{D}$ is true (T).

Therefore the truth-table for $\mathcal{D}$ is identical with that of $f(x_1,x_2,x_3)$.

Example 1.3.4 Consider the truth-function $f(x_1,x_2,x_3)$ given by the truth-table:

x_1	x_2	x_3	$f(x_1,x_2,x_3)$
T	T	T	F
T	F	T	F
T	T	F	F
T	F	F	F
F	T	T	F
F	F	T	F
F	T	F	F
F	F	F	F

Note that there is no row in which the last entry is T, i.e. all the last entries are F. In such a case we construct a statement form $\mathcal{D}$ as follows:-

$$\mathcal{D} = (A_1 \wedge \sim A_1) \vee (A_2 \wedge \sim A_2) \vee (A_3 \wedge \sim A_3)$$

Then $\mathcal{D}$ generates $f(x_1,x_2,x_3)$. For this we have only to check that $\mathcal{D}$ takes the truth-value F for any truth-assignment of A_1, A_2, A_3. But this is almost obvios, because for any assignment of truth-value one of A_i and $\sim A_i$ (i = 1,2,3) is true and the other false, showing that all of $A_1 \wedge \sim A_1$, $A_2 \wedge \sim A_2$, $A_3 \wedge \sim A_3$ are false, and therefore $\mathcal{D}$ is false. Thus, the truth-table for $\mathcal{D}$ is identical with that of $f(x_1,x_2,x_3)$.

The above examples explian the steps in the proof of a general result on the adequacy of three connectives, presented in the next proposition.

Proposition 1.3.1 Every truth-function can be generated by a statement form involving only the three connectives $\sim$, $\wedge$ and $\vee$.

Proof: Let $f(x_1,x_2,....,x_n)$ be a truth-function in n variables. Our problem is to construct a statement form $\mathcal{D}$ in n statement letters A_1, A_2,...., A_n such that the truth-tables for $\mathcal{D}$ and f are identical.

The truth-table for f has 2^n rows of entries let R_i denote the i-th row of entries. The last entry in a row is the truth-value of the function f corresponding to the truth-values of $x_1, x_2, ..., x_n$ in the same row.

Case I Suppose all the last entries are F

Construct statement form $\mathcal{D}$ as follows:

$\mathcal{D} = (A_1 \wedge \sim A_1) \vee (A_2 \wedge \sim A_2) \vee ... \vee (A_n \wedge \sim A_n)$. For any assignment of truth-values to $A_1, A_2, ..., A_n$ all of $A_1 \wedge \sim A_1, \vee (A_2 \wedge \sim A_2, ..., A_n \wedge \sim A_n$ are false (F), and therefore $\mathcal{D}$ is false (F). Hence the truth-table for $\mathcal{D}$ is identical with that of f, i.e. $\mathcal{D}$ generates f.

Case II. Suppose not all the last entries are F.

Pick up those rows $R_{j_1}, R_{j_2}, ..., R_{j_k}$ for which the last entries are T. Now for such a row R_{j_k} construct the conjunct

$\mathcal{D}_{j_k} = A_{1 j_k} \wedge A_{2 . j_k} \wedge ... \wedge A_{n j_k}$, where

$A_{i j_k} = A_i$ or $\sim A_i$

according as the truth-value in the ij_k-th entryis T or F.

Then construct

$$\mathcal{D} = \mathcal{D}_{j_1} \vee \mathcal{D}_{j_2} \vee .... \vee \mathcal{D}_{j_k}$$

W claim that $\mathcal{D}$ generates $f(x_1, x_2, ..., x_n)$.

This follows from the following observations:-

(1) If we give to $A_1, A_2, ..., A_n$ an assignment of truth-values same as that in a row R_i where $i \neq j_1, j_2, ..., j_k$, then all of $\mathcal{D}_{j_1}, \mathcal{D}_{j_2}, ... \mathcal{D}_{j_k}$ are false (F), and therefore $\mathcal{D}$ is false (F).

(2) On the other hand if we give to $A_1, A_2, ..., A_n$ an assignment of truth-values same as that in a row R_{j_k}, then D_{j_k} is true (T). This shows that the truth-table for Dis same as that for f. Hence $\mathcal{D}$ generates f.

Corollary : Every truth-function can be generated by a statement form containing as connectives

(i) only $\wedge$ and $\sim$

or (ii) only $\vee$ and $\sim$

or (iii) only $\rightarrow$ and $\sim$

[This means that the number of adequate connectives may even be reduced to two.]

Proof: (i) It can be checked that for any two statement letters A and B.

$$A \vee B \leftrightarrow (\sim A \wedge \sim B)$$

is a tautology, i.e. $A \vee B$ and $\sim A \wedge \sim B$ are logically equivalent, i.e. truth-tables for $A \vee B$ and $\sim A \wedge \sim B$ are identical, i.e. the connective $\vee$ can be replaced by the connectives $\wedge$ and $\sim$. Therefore $\wedge$ and $\sim$ can serve as an adequate set of connectives in place of $\wedge$, $\vee$ and $\sim$.

(ii) It can be checked that

$$A \wedge B \leftrightarrow \sim(\sim A \vee \sim B) \text{ is a tautology.}$$

So can be replaced by $\vee$ and $\sim$.

(iii) It can be checked that

$$A \vee B \leftrightarrow \sim A \to B$$

$$A \wedge B \leftrightarrow \sim(A \to \sim B)$$

are tautologies. Therefore $\vee$ can be replaced by $\to$ and $\sim$, and $\wedge$ can also be replaced by $\to$ and $\sim$.

Definition : A disjunctive normal form is a statement form which is a disjunction such that each disjunct is a conjunction of statement letters and negations of statement letters.

For example

$$(\sim A_1 \wedge A_2) \vee (A_1 \wedge \sim A_2) \vee (A_1 \wedge \sim A_2)$$

is a disjunctive normal form.

With this definition, the statement form $\mathcal{D}$ of Proposition 1.3.1 is a disjunctive normal form, and therefore the proposition can be restated as:

Every truth-function can be generated by a disjunctive normal form.

Proposition 1.3.2 Every statement form is logically equivalent to a disjunctive normal form.

Proof : Let $\mathcal{A}$ be any statement form in n statement letters $A_1, A_2, \ldots, A_n$. Suppose $\mathcal{A}$ gives rise to the truth-function $f(x_1, x_2, \ldots, x_n)$. Then $\mathcal{A}$ and f have identical truth-tables. Suppose f is generated by the statement form $\mathcal{D}$. Then the truth-tables of f and $\mathcal{D}$ are identical, and, as mentioned just before this proposition, $\mathcal{D}$ is a disjunctive normal form. Thus the truth-tables of $\mathcal{A}$ and the disjunctive normal form $\mathcal{D}$ are identical; but this is same thing to say that $\mathcal{A} \leftrightarrow \mathcal{D}$ is a tautology, i.e. $\mathcal{A}$ and $\mathcal{D}$ are logically equivalent.

Definition : A conjunctive normal form is a statement form which is a conjunction such that each conjunct is a disjunct of statement letters and negations of statement letters.

For examples

$$(\sim A_1 \vee A_2) \wedge (A_1 \vee \sim A_2) \wedge (\sim A_1 \vee \sim A_2)$$

is a conjunctive normal form.

Proposition 1.3.3 Any statement form is logically equivalent to a conjunctive normal form.

Prof: Let $\mathcal{A}$ be any statement form. Then $\sim\mathcal{A}$ is also a statement form. So $\sim\mathcal{A}$ is logically equivalent to a disjunctive normal form $\mathcal{D}$, i.e. a statement form which is a disjunction such that each disjunct is a conjunction of statement letters and negations of statement letters. Therefore $\mathcal{A}$ is logically equivalent to $\sim\mathcal{D}$. Now by repeated application of the tautologies (check!)

$$\sim(\sim A) \leftrightarrow A$$

$$\sim(A \vee B) \leftrightarrow \sim A \wedge \sim B$$

$$\text{and } \sim(A \wedge B) \leftrightarrow \sim A \vee \sim B$$

we see that $\sim\mathcal{D}$ is logically equivalent to a statement form which is a conjunction such that each conjunct is a disjunction of statement letters and negations of statement letters, i.e. $\mathcal{A}$ is logically equivalent to a conjunctive normal form.

1.4 Two Other Binary Connectives : Joint Denial and Alternative Denial

We have so far introduced four **binary** connectives, namely $\wedge$, $\vee$, $\rightarrow$ and $\leftrightarrow$. These binary connectives are **completely determined** by their truth-tables :-

A	B	A∧B
T	T	T
F	T	F
T	F	F
F	F	F

Table 1

A	B	A∨B
T	T	T
F	T	T
T	F	T
F	F	F

Table 2

A	B	A→B
T	T	T
F	T	T
T	F	F
F	F	T

Table 3

A	B	A↔B
T	T	T
F	T	F
T	F	F
F	F	T

Table 4

The order of the entries in the last column of each of the above tables determines the corresponding binary connective. There are altogether 16 ways of constructing the last column in a truth-table of the form given above, and each way gives rise to a binary connective b between two statement letters A and B i.e. each way gives the truth-value of AbB corresponding to a given assignment of truth-values to A and B. Tables 1, 2, 3, 4 define AbB as $A\wedge B$, $A\vee B$, $A\rightarrow B$, $A\leftrightarrow B$ respectively. We now consider two other binary connectives (determined by truth tables) which are important in many respects :-

The binary connective ↓, called **joint denial**, is defined by the truth-table.

A	B	A↓B
T	T	F
F	T	F
T	F	F
F	F	T

or in the concise form

A	↓	B
T	F	T
F	F	T
T	F	F
F	T	F
1	2	1

Note that $A \downarrow B$ is true when and only when both A and B are false, i.e. neither A nor B is true. The binary connective $\mid$, called **alternative denial** is defined by the truth-table.

A	B	$A \mid B$
T	T	F
F	T	T
T	F	T
F	F	T
1	2	1

or in the concise form.

A	$\mid$	B
T	F	T
F	T	T
T	T	F
F	T	F

Proposition 1.4.1 Every truth-function can be generated by a statement form containing a single binary connective $\downarrow$ (jont denial) or $\mid$ (alternative denial).

Proof : We know that (Corollary to Proposition 1.3.1) every truth-function can be generated by a statement form containing as connectives $\wedge$ and ~ only. We show that both $\wedge$ and ~ can be 'replaced' by $\downarrow$ alone, i.e. we show that the following two are tautologies :

$$A \wedge B \leftrightarrow ((A \downarrow A) \downarrow (B \downarrow B))$$

$$\sim A \leftrightarrow (A \downarrow A)$$

That the above two are tautologies are seen by constructing a truth-table for each of them, and in such a construction we use the truth-table for $\downarrow$:

A	$\wedge$	B	$\leftrightarrow$	((A	$\downarrow$	A)	$\downarrow$	(B	$\downarrow$	B))
T	T	T	T	T	F	T	T	T	F	T
F	F	T	T	F	T	F	F	T	F	T
T	F	F	T	T	F	F	F	F	T	F
F	F	F	T	F	T	F	F	F	T	F
1	2	1	4	1	2	1	3	1	2	1

~	A	$\leftrightarrow$	)A	$\downarrow$	A_
F	T	T	T	F	T
T	F	T	F	T	F
2	1	3	1	2	1

So $\wedge$ and $\sim$ can be replaced by $\downarrow$ alone, and this shows that a truth-function canbe generated by a statement form containing the single binary connectives $\downarrow$.

From Corollary to Proposition 1.3.1 we also know that every truth-function can be generated by a statement form containing as connectives $\vee$ and $\sim$ only. Now we show that $\vee$ and $\sim$ can be 'replaced by' by $\mid$ alone, and for this it will be sufficient to show that the following two are tautologies :

$$A \vee B \leftrightarrow ((A \mid A) \mid (B \mid B))$$

$$\sim A \leftrightarrow (A \mid A)$$

We now construct the truth-tables by using the truth-table for $\mid$:

A	$\vee$	B	$\leftrightarrow$	((A	$\mid$	A)	$\mid$	(B	$\mid$	B))
T	T	T	T	T	F	T	T	T	F	T
F	T	T	T	F	T	F	T	T	F	T
T	T	F	T	T	F	T	T	F	T	F
F	F	F	T	F	T	F	F	F	T	F
1	2	1	4	1	2	1	3	1	2	1

$\sim$	A	$\leftrightarrow$	(A	$\mid$	A)
F	T	T	T	F	T
T	F	T	F	T	F
2	1	3	1	2	1

This proves the proposition.

We now ask a very important question-

Does there exist any **other** binary connective that alone can be adequate to construct a statement form which generates a given truth-function ? The answer is No as will be seen from the next proposition.

Proposition 1.4.2 The only binary connectives that alone are adequate for generation of all truth-functions are $\downarrow$ and $\mid$.

Proof : Let b be a connective which alone is adequate for generation of all truth-functions. It will be sufficient to show that AbB is either $A \downarrow B$ or $A \mid B$, i.e. the truth-table for AbB is

either

A	b	B
T	F	T
F	F	T
T	F	F
F	T	F

or

A	b	B
T	F	T
F	T	T
T	T	F
F	T	F

Suppose in the truth-table for AbB, AbB takes the value T when both A and B take the value T. Then **any** statement form built up by using the only connective b would take the value T when all the statement letters take the value T. But the statement form ~A does not take the value T when its statement letter A takes the value T. This gives us a contradiction that AbB takes the value T when both A and B take the value T. So AbB must take the value F when both A and B take the value T.

Next, suppose AbB takes the value F when both A and B takes the value F. Then any statement form built up by using the only connective b would take the value F when all the statement letters take the value F. But we have the statement form ~A which takes the value T when A takes the value F. This gives a contradiction, and so AbB must take the value T when both A and B take the value F. Thus we have the following partial table for AbB.

A	b	B
T	F	T
F	*	T
T	**	F
F	T	F

If the values in the positions * and ** in the above are F, F or T, T, then p is ↓ or |. We prove that the values in these positions **cannot** be F, T or T, F. If the values are F, T then (A b B) ↔ ~B is a tautology which can be seen from the following table :

(A	b	B)	↔	~	B
T	F	T	T	F	T
F	F	T	T	F	T
T	T	F	T	T	F
F	T	F	T	T	F
1	2	1	3	2	1

This means that b can be replaced by ~, i.e. ~alone is also adequate for generation of all truth-functions. But we have the statement for mAB which cannot be expressed by using ~only. Therefore, in the positions *, ** we cannot have F, T. If in the positions *, ** we have the values T, F then (AbB) ↔ ~A is a tautology as seen from the following table :

(A	b	B)	↔	~	B
T	F	T	T	F	T
F	T	T	T	T	F
T	F	F	T	F	T
F	T	F	T	T	F
1	2	1	3	2	1

This means that b can be replaced by ~, i.e. ~alone is adequate for generation of all truth-functions. But this is not possible as we have seen above. Therefore the two positions *, ** must be filled up by F, F or T, T, i.e. the truth-table for AbB is

either

A	b	B
T	F	T
F	F	T
T	F	F
F	T	F

or

A	b	B
T	F	T
F	T	T
T	T	F
F	T	F

i.e. b is either ↓ or |. This proves the proposition.

1.5 Arguments and Validity

Our intuitive notion of an argument is that we make some assertions and then we come to a conclusion. In other words, an argument is a collection of assertions followed by a conclusion. The assertions in an argument are referred to as the **premisses** of the argument. In day to day life very often we have to 'deal with' arguments, and analyse if an argument put forward by some one **makes any sense** or **is valid.**

These things will now be discussed in the light of mathematical logic.

Let us consider the following examples of arguments.

1. (a) If **Prasanta is a man,** then **Prasanta is mortal**
   (b) Therefore **Prasanta is mortal**

2. (a) **Prasanta is a man**
   (b) Therefore **Prasanta is mortal**

3. (a) If **the function f is not continuous,** then **the function g is not differentiable**

(b) **The function g is differentiable**

(c) Therefore the **function f is continuous**

4. (a) If **Prasanta has owned a house,** then **either he has sold his car or he has borrowed money from the bank.**

(b) **Prasanta has not borrowed money from the bank.**

(c) Therefore, if **Prasanta has not sold his car,** then **he has not owned a house.**

The above arguments can be symbolised by using statement letters for the atomic statements and appropriate connectives.

(1) Using A for **Prasanta is a man,**
B for **Prasanta is mortal,**
the argument can be symbolised as
(a) $A \rightarrow B$
(b) A
(c) $\therefore$ B
$A \rightarrow B$ and A are the premisses, and B is the conclusion.

(2) Using A for **Prasanta is a man,**
B for **Prasanta is mortal,**
the argument can be symbolised as
(a) A
(b) $\therefore$ B
A is the only premiss, and B is the conclusion.

(3) Using A for **the function f is continuous,**
B for **the function g is differentiable,**
the argument can be symbolised as
(a) $\sim A \rightarrow \sim B$
(b) B
(c) $\therefore$ A
$\sim A \rightarrow \sim B$, B are the premisses, and A is the conclusion.

(4) Using A for **Prasanta has owned a house,**
B for **Prasanta has sold his car,**
C for **Prasanta has borrowed money from the bank,**

the argument can be symbolised as

(a) $A \to B \vee C$

(b) $\sim C$

(c) $\therefore \sim B \to \sim A$

Here $A \to B \vee C$ and $\sim C$ are the premisses, and $\sim B \to \sim A$ is the conclusion.

From the above examples it becomes clear that when symbolised, an argument takes the form of a sequence of statement forms

$\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n, \therefore \mathcal{A}$

in which $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n$ are the premisses and $\mathcal{A}$ is the conclusion. Such a sequence symbolising an argument is called an **argument form.**

Validity of an Argument :

Let $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n, \therefore \mathcal{A}$ be an argument form symbolising an argument. $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n$ and $\mathcal{A}$ are statement forms in some statement letters. For a given assignment of truth-values to the statement letters we can determine the truth-values of $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n$ and $\mathcal{A}$ by constructing truth-tables. For some assignments of truth-values to the statement letters all of $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n$; may take the truth-value T, but then $\mathcal{A}$ may take the value either T or F.

If $\mathcal{A}$ **always** takes the value T whenever all of $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n$ take the value T for an assignment of truth-values to the statement letters, then the argument form, and the corresponding argument, are said to be **valid.**

Let us now examine the validity of the arguments given in Example 1-4 above :

1. The argument form symbolising the given argument is

   $A \to B, A; \therefore B$

   We construct a **combined** truth-table for $A \to B$, A and B as follows :

Premisses			Conclusion	
A	→	B	A	B
T	T	T	T	T
F	T	T	F	T
T	F	F	T	F
F	F	F	F	F

Only in the first row both the premisses take the value T, and the table shows that the corresponding value of the conclusion is also T. So the argument form, and therefore the argument, are valid.

2. The argument form symbolising the given argument is

   $A; \therefore B$

The **combined** truth-table for A and B is as follows :

A	B
Ⓣ	T
F	T
Ⓣ	F
F	F

Hence the premiss A take the value T in the first as well as the third row, but the conclusion B takes the value T in the first row and the F in the third row. Therefore the argument is valid.

3. The argument form symbolising the given argument is
   ~A →~B, B; ∴ A
   The **combined** truth-table is :

~	A	→	~	B	B	A
F	T	Ⓣ	F	T	Ⓣ	T
T	F	F	F	T	T	F
F	T	T	T	F	F	T
T	F	T	T	F	F	F

Only in the first row both the premisses take the value T, and the corresponding value of the conclusion is also T. This means that the argument form, and therefore the argument, are valid.

4. Here the argument form is
   A → B v C, ~C; ∴ ~B → ~A
   The **combined** truth-table is :

A	→	B	v	C	~	C	~	B	→	~	A
T	T	T	T	T	F	T	F	T	T	F	T
T	T	F	T	T	F	T	T	F	T	F	T
T	Ⓣ	T	T	F	Ⓣ	F	F	T	T	F	T
T	F	F	F	F	T	F	T	F	F	F	T
F	T	T	T	T	F	T	F	T	T	T	F
F	T	F	T	T	F	T	T	F	T	F	F
F	Ⓣ	T	T	F	Ⓣ	F	F	T	T	T	F
F	Ⓣ	F	F	F	Ⓣ	F	T	F	T	T	F

Both the premisses take the value T in the third, seventh and eight rows, and the corresponding value of the conclusion in his rows is also T. Therefore the argument form, and therefore the argument, are valid.

The above discussion and examples suggest the following useful proposition.

Proposition 1.5.1 The argument form

$\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n; \therefore \mathcal{A}$ is valid if and only if the statement form $\mathcal{A}_1 \wedge \mathcal{A}_2 \wedge \ldots \wedge \mathcal{A}_n \rightarrow \mathcal{A}$ is a tautology.

Proof : Suppose first

$\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n; \therefore \mathcal{A}$ is a valid argument form, but $\mathcal{A}_1 \wedge \mathcal{A}_2 \wedge \ldots \wedge \mathcal{A}_n \rightarrow \mathcal{A}$ is not a tautology. Then there is an assignment of truth-values of the statement letters for which $\mathcal{A}_1 \wedge \mathcal{A}_2 \wedge \ldots \wedge \mathcal{A}_n$ is TRUE but $\mathcal{A}$ is FALSE.

If $\mathcal{A}_1 \wedge \mathcal{A}_2 \wedge \ldots \wedge \mathcal{A}_n$ is true, then all of $\mathcal{A}_1 \wedge \mathcal{A}_2 \wedge \ldots \wedge \mathcal{A}_n$ are true. It contradicts that the argument form is valid (when all the premisses are true, then the conclusion must be true).

Conversely suppose $\mathcal{A}_1 \wedge \mathcal{A}_2 \wedge \ldots \wedge \mathcal{A}_n \rightarrow \mathcal{A}$ is a tautology, but the argument form $\mathcal{A}_1, \mathcal{A}_2, \ldots \mathcal{A}_n; \therefore \mathcal{A}$ **is not** valid. Then if all of $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n$ take the value T. $\mathcal{A}$ must take the value of F which gives a contradiction since $\mathcal{A}_1 \wedge \mathcal{A}_2 \wedge \ldots \wedge \mathcal{A}_n \rightarrow \mathcal{A}$ is a tautology.

Example 1.5.1 Examine the validity of the following arguments :

(i) $p_1 \rightarrow (p_2 \rightarrow p_3), p_2; \therefore p_1 \rightarrow p_3$

(ii) If there is a gold mine in Kamakhya hill, then either the experts are right or the government is lying. There is no gold mine in Kamakhya hill or the experts are wrong. Therefore the government is not lying.

(i) The argument is already in symbolic form. We have to check, in view of

Proposition 1.5.1 : If $(p_1 \rightarrow (p_2 \rightarrow p_3)) \wedge p_2 \rightarrow (p_1 \rightarrow p_3)$ is a tautology. So we construct the following truth-table :

$(p_1$	$\rightarrow$	$(p_2$	$\rightarrow$	$p_3))$	$\wedge$	p_2	$\rightarrow$	$(p_1$	$\rightarrow$	$p_3)$
T	T	T	T	T	T	T	T	T	T	T
T	T	F	T	T	F	F	T	T	T	T
T	F	T	F	F	F	T	T	T	F	F
T	T	F	T	F	F	F	T	T	F	F
F	T	T	T	T	T	T	T	F	T	T
F	T	F	T	T	F	F	T	F	T	T
F	T	T	F	F	T	T	T	F	T	F
F	T	F	T	F	F	F	T	F	T	F
1	3	1	2	1	4	1	5	1	2	1

The above table gives a tautology. Therefore the given argument form is valid

(ii) First we have to symbolise the given argument. Let us use

A **for** There is a gold mine in Kamakhya hill,

B **for** The experts are right,

C **for** The government is lying

We now have the following arguments form :-

$A \to (B \vee C)$, $\sim A \vee B$, $\therefore \sim C$

Then we use Proposition 1.5.1 :-

(A	→	(B	∨	C))	∧	(~	A	∨	~	B)	→	~	C
T	T	T	T	T	F	F	T	F	F	T	T	F	T
T	T	F	T	T	T	F	T	T	T	F	F	F	T
T	T	T	T	F	F	F	T	F	F	T	T	T	F
T	F	F	F	F	F	F	T	T	T	F	T	T	F
F	T	T	T	T	T	T	F	T	F	T	F	F	T
F	T	F	T	T	T	T	F	T	T	F	F	F	T
F	T	T	T	F	T	T	F	T	F	T	T	T	F
F	T	F	F	F	T	T	F	T	T	F	T	T	F
1	4	1	3	1	5	2	1	3	2	1	6	2	1

The above table does not give a tautology. Therefore the given arguent is **not** valid.

Summary :

- If $\mathcal{A} \to \mathcal{B}$ is a tautology, then we say that $\mathcal{A}$ logically implies $\mathcal{B}$ or that $\mathcal{B}$ is a logical consequence of $\mathcal{A}$.
  If $\mathcal{A} \leftrightarrow \mathcal{B}$ is a tautology then $\mathcal{A}$ and $\mathcal{B}$ are said to be logically equivalent and written as $\mathcal{A} \equiv \mathcal{B}$.
- If $\mathcal{A}$ and $\mathcal{B}$ are tautologies then so is $\mathcal{B}$.
- Every truth-function can be generated by statement form involving only the three connectives $\sim$, $\wedge$ and $\wedge$.
- Every statement form is logically equivalent to disjunctive normal form.
- Any statement form is logically equivalent to conjunctive normal form.
- The argument form $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n$; $\therefore \mathcal{A}$ is valid if and only if the statement-form $\mathcal{A}_1 \wedge \mathcal{A}_2 \wedge \ldots \wedge \mathcal{A}_n \to \mathcal{A}$ is a tautology.

G.U. Questions

1996

1. Translate the following sentences into symbolic notation using statement letters for the atomic sentences and the usual connectives :

   (i) If demand has remained constant and prices have been increased, then turnover must be decreased.

   (ii) If y is an integer then z is not real, provided that x is a rational number. 1+1=2

2. Give examples with justification of tautology, contradiction and contingent containing at least two statement letters. 2+2+2=6

3. Write the following statement in symbolic form

'If either labour or management is stubborn, then the strike wil be settled if and only if the government obtains an injunction, but troops are not sent into the mills'

Now determine by a truth value analysis whether the above statement is **true** or **false** under the following assumptions:

'If the government obtains an injunction, then the troops will be sent into the mills. If troops are sent into the mills, then the strike will be settled Management is stubborn'. 8

4. Define the binary connectives **joint denial** (↓) and **alternative denial** (|). Prove that ↓ and | are the only binary connectives that alone are adequate for representation of all truth functions. 2+6=8

1997

1. Translate the following sentences into symbolic notation using statement letters for the atomic sentences and usual connectives :

   (i) If the murderer has not left the country, then somebody is harbouring him.

   (ii) If Das is not elected leader of the party, then either Chaliha or Saikia will leave the Cabinet, and we shall lose the election.

   (iii) The sum of two numbers is even if and only if either both numbers are even or both numbers are odd. 1+1+1=3

2. When are two statement forms A and B said to be **logically equivalent**? Is $\sim(A \wedge B)$ logically equivalent to $\sim A \wedge \sim B$ where A and B are two statement letters? 1+2=3

3. A and B are two statement forms and A and $A \rightarrow B$ are both tautologies. Prove that B is a tautology.

4. Suppose A is a statement form which contains another statement form A_1. Suppose B is obtained from A by substituting B_1 for one or more occurences of A_1. Then prove that $A_1 \rightarrow B_1$ logically implies $A \leftrightarrow B$. 5

5. Let $f(x_1,x_2)$ be a truth function with the following truth table :

x_1	x_2	$f(x_1,x_2)$
T	T	F
F	T	T
T	F	T
F	F	T

Construct (with justification) a statement from D with the three connectives ~, $\wedge$ and $\vee$ only such that D generates $f(x_1,x_2)$. 6

1998

1. Translate the following sentences into symbolic forms using statement letters for the atomic sentences and the usual connectives :
   (i) He is tall, or he is short and handsome.
   (ii) To be poor but happy means to be rich. 1+2=3
2. Examine the validity of the following argument by writing the corresponding argument form :
   'If the function f is not continuous, then the function g is not differentiable. The function g is differentiable. Therefore, the function f is continuous. 5
3. When is a statement form said to be a **tautology?** Examine if the following statement form is a tautology.
   $((\sim p)\rightarrow q)\rightarrow(p\rightarrow(\sim q))$ 1+2=3
4. Prove that every truth function is generated by a statement form involving the connectives , , and only. Show further that the adequate number of connectives may even be reduced to two : $\wedge$ and ~, or and ~, $\vee$ or $\rightarrow$ and ~. 5+2+2+2=11
5. Defind **disjunctive normal form** and **conjunctive normal form**. Show that any statement form is logically equivalent to a conjunctive normal form. 1+1+3=5

1999

1. Translate the following composite sentences into symbolic form using to stand for the atomic sentences :
   (i) If demand has remained constant and prices have been increased, then turnover must be decreased.
   (ii) If Jones is not elected leader of the party, then either Smith or Robinson will leave the cabinet, and we shall lose the election.
   (iii) The corps will survive if and only if irrigation ditches are dug; should be crops not survive, then the farmers will go bankrupt and leave. 2+2+2=6
2. Define **tautology, contradiction and contingent** and give an example (with justification) of each. 2+2+2=6
3. Construct (with justification) a statement form that generates the truth function $f(x_1,x_2)$ define by the following truth table : 4

x_1	x_2	$f(x_1,x_2)$
T	T	F
F	T	T
T	F	T
F	F	T

4. Define the binary connectives **joint denial** (↓) and **alternative denial** (|) of propositional calculus. Prove that the only binary connectives that alone are adequate for representation of all truth functions are and .
   2+6=8

Unit-2
Formal Statement Calculus

Introduction :

The word 'formal' is one which appears regularly in logic text books without being explained. It is used when referring to a situation where symbols are being used and where the behaviour and properties of the symbols are determined completely by a given set of rules. In a formal system the symbols have no meanings, and in dealing with them we must be careful to assume nothing about their properties other than what is specified in the system. .

2.1. Notion of a Formal Theory :

A **formal theory** $\mathcal{E}$ is said to be **defined** if the following conditions are satisfied:

(i) There is a countable set of symbols given as the symbols of $\mathcal{E}$. A finite sequence of symbols of $\mathcal{E}$ is called an **expression** of $\mathcal{E}$.

(ii) There is a subset of the set of all expressions of $\mathcal{E}$, called the set of **well-formed formulas,** abbreviated wffs or simply wfs, of $\mathcal{E}$. There is usually an effective procedure to determine whether a given expression is a wf or not.

(iii) A subset of the set of all wfs is set aside and is called a **set of axioms.** There is usually an effective way to determine whether a given wf is an axiom or not. In such a case $\mathcal{E}$ is called an **Axiomatic Theory.**

(iv) There is a finite set $R_1, R_2,...,R_n$ of relations among the wfs called the **rules of inference.** For each R_i there is a unique positive integer j such that for every set of j wfs and any given wf $\mathcal{A}$ one can decide whether the given j wfs are in the relation R_i to $\mathcal{A}$. If it is so then $\mathcal{A}$ is called **a direct consequence of the given j wfs by virtue of the R_i.**

Some common terms like **proof, theory, consequence, deduction** are given precise meanings in a formal theory $\mathcal{E}$ as follows:

(a) A **proof** in $\mathcal{E}$ is a sequence $\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_n$ of wfs such that for each i, either $\mathcal{A}_i$ is an axiom of $\mathcal{E}$ or $\mathcal{A}_i$ is a direct consequence of some of the preceding wfs by virtue of one of the rules of inference.

(b) A **theorem** in $\mathcal{E}$ is a wf $\mathcal{A}$ such that there is a proof $\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_n$ where $\mathcal{A}_n$ is same as $\mathcal{A}$. Such a proof is called a **proof of $\mathcal{A}$.**

(c) A wf $\mathcal{A}$ is said to be a **consequence** in $\mathcal{E}$ of a set Γ of wfs if and only if there is a sequence $\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_n$ of wfs such that $\mathcal{A} = \mathcal{A}_n$ and for each i, **either**

(i) $\mathcal{A}_i$ is an axiom

or (ii) $\mathcal{A}_i$ is a direct consequence by some rule of inference of some of the preceding wfs in the sequence $\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_n$.

or (iii) $\mathcal{A}_i$ is in Γ.

Such a sequence $\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_n$ is called a **deduction of $\mathcal{A}$ from Γ**(or a proof of $\mathcal{A}$ from Γ). The members of Γ are called the **hypotheses** or the **premisses** of the deduction.

If $\mathcal{A}$ is a consequence in a formal theory $\mathcal{E}$ of the set of premisses Γ, then we write in symbol as

$$\Gamma \underset{\mathcal{E}}{\vdash} \mathcal{A}$$

or simply $\Gamma \vdash \mathcal{A}$ if there is no need to specify the formal theory $\mathcal{E}$.

If Γ is a finite set $\{\mathcal{B}_1, \mathcal{B}_2, ..., \mathcal{B}_n\}$, instead of writing

$$\{\mathcal{B}_1, \mathcal{B}_2, ..., \mathcal{B}_n\} \vdash \mathcal{A}$$

we usually write

$$\mathcal{B}_1, \mathcal{B}_2, ..., \mathcal{B}_n \vdash \mathcal{A}$$

It may happen that $\Gamma = \phi$, the empty set. From the above definition of consequence it is clear that $\phi \vdash \mathcal{A}$ if and only if $\mathcal{A}$ is a theorem. We therefore use the symbol $\vdash \mathcal{A}$ (which is **logically** same as $\phi \vdash \mathcal{A}$) to mean that $\mathcal{A}$ is a theorem.

The following three properties of consequence follow from the definition of consequence:

(1) If $\Gamma \subseteq \Delta$ and $\Gamma \vdash \mathcal{A}$, then $\Delta \vdash \mathcal{A}$

[This means that if $\mathcal{A}$ is **provable** from a set Γ of premisses, then $\mathcal{A}$ is also provable from a larger set of premisses.]

(2) $\Gamma \vdash \mathcal{A}$ if and only if there is a finite subset Δ of Γ such that $\Delta \vdash \mathcal{A}$

[Half of this follows from (1): If $\Delta \subseteq \Gamma$ and $\Delta \vdash \mathcal{A}$ then from (1) $\Gamma \rightarrow \mathcal{A}$. The other half follows from the definition of a proof i.e. in a proof there are a finite number of premisses.]

(3) If $\Delta \vdash \mathcal{A}$, and for each $\mathcal{B} \in \Delta$, $\Gamma \vdash \mathcal{B}$ then $\Gamma \vdash \mathcal{A}$

[This means that if $\mathcal{A}$ is provable from premisses in Δ, and each premiss in Δ is provable from the premisses in Γ, then $\mathcal{A}$ is provable from premisses in Γ which is quite obvious.]

2.2. Propositional Calculus as Formal(Axiomatic) Theory:

We recall that in order to have a formal(axiomatic) theory we must have the following:

1. A countable set of symbols.
2. A set of well-formed formulas(wfs).
3. A set of axioms.
4. A set of rules of inference.

A formal(axiomatic) theory L of propositional calculus may be introduced by prescribing the above four things as follows:

1. $\sim, \rightarrow, (...),$, and the statement setters $A_1, A_2, ...$ are taken as the symbols for L.

2. (a) All statement letters are taken as wfs.

(b) If $\mathcal{A}$ and $\mathcal{B}$ are wfs by (a), then $\sim \mathcal{A}$ and $\mathcal{A} \rightarrow \mathcal{B}$ are also taken as wfs.

(c) A statement form that is **generated** by wfs of types (a) and (b) is also a wf.

3. If $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ are any wfs in L, then the following are taken as axioms:

(A1) $\mathcal{A} \to (\mathcal{B} \to \mathcal{A})$

(A2) $(\mathcal{A} \to (\mathcal{B} \to \mathcal{C}) \to ((\mathcal{A} \to \mathcal{B}) \to (\mathcal{A} \to \mathcal{C}))$

(A3) $(\sim \mathcal{A} \to \sim \mathcal{B}) \to (\mathcal{B} \to \mathcal{A})$

The above axioms are usually called **axiom schemas,** rather than simply axioms, because each one of them actually gives an infinite number of axioms being valid for **any set** of three wfs $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$. Note that **schema** means **plan** or **diagram.** It is also important to bear in mind that the choice of axioms is not unique-this is only one choice out of many probables.

4. The only rule of inference is: A wf $\mathcal{B}$ is the direct consequence of the wfs $\mathcal{A}$ and $\mathcal{A} \to \mathcal{B}$.

This rule of inference is usually called the **Modus Ponens,** abbrevialed M.P., which means in Latin a **mode of expression**.

It is worthwhile to note that all the axioms (A1), (A2), (A3) can easily be verified to be tautologies in the sense of informal propositional calculus(Unit I) and the M.P. is suggested by a proposition proved in Unit I which states that if $\mathcal{A}$ and $\mathcal{A} \to \mathcal{B}$ are tautologies then so is $\mathcal{B}$.

Now the definitions of **proof, theorem and deduction** given in § 2.1. can be restated in the context of the theory L as follows:

A **proof** is a sequence of wfs $\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_n$ such that each $\mathcal{A}_i$ is an axiom or is deducible from some of the preceding wfs and the M.P.

A **theorem** $\mathcal{A}$ is a wf such that there is a proof $\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_n$ with $\mathcal{A} = \mathcal{A}_n$ (notation : $\underset{L}{\vdash} \mathcal{A}$)

A **deduction** $\mathcal{A}$ from a set Γ of wfs is a sequence $\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_n$ of wfs such that $\mathcal{A} = \mathcal{A}_n$ where each $\mathcal{A}_i$ is

(i) an axiom

or (ii) a wf in Γ

or (iii) deducible from some of the preceding wfs and the M.P.(notation : $\Gamma \underset{L}{\vdash} \mathcal{A}$)

Example 2.2.1.: For all wfs $\mathcal{A}$ and $\mathcal{B}$ of L

$$\underset{L}{\vdash} (\mathcal{A} \to \mathcal{B}) \to (\mathcal{A} \to \mathcal{A})$$

i.e. $(\mathcal{A} \to \mathcal{B}) \to (\mathcal{A} \to \mathcal{A})$ is a theorem of L.

Here we have to show that

$$(\mathcal{A} \to \mathcal{B}) \to (\mathcal{A} \to \mathcal{A})$$

is provable by repeated applications of the axioms and the M.P. We have the following sequence of wfs:

1. $\mathcal{A} \to (\mathcal{B} \to \mathcal{A})$ (by (A1)

2. $(\mathcal{A}\to(\mathcal{B}\to\mathcal{A}))\to((\mathcal{A}\to\mathcal{B})\to(\mathcal{A}\to\mathcal{A}))$ by (A2)

[writing $\mathcal{A}$ for $\mathcal{C}$ in (A2)]

3. $(\mathcal{A}\to\mathcal{B})\to(\mathcal{A}\to\mathcal{A})$ by (1), (2) and M.P.

The last member of the above sequence is same as the wf given to establish as a theorem.

The above example illustrates the **proof** of a **theorem**. We have a sequence of wfs (1, 2 and 3) such that each one of them is either an axiom or a direct consequence of some of the preceding wfs in the sequence by virtue of the rule of inference(M.P.).

Here 1 and 2 are axioms and 3 is the direct consequence of 1 and 2 by virtue of the M.P.

Example 2.2.2.: For any two wfs $\mathcal{A}$ and $\mathcal{B}$ of L

$$\vdash_{L} \sim\mathcal{B}\to(\mathcal{B}\to\mathcal{A})$$

i.e. $\sim\mathcal{B}\to(\mathcal{B}\to\mathcal{A})$ is a theorem of L.

We have to show that $\sim\mathcal{B}\to(\mathcal{B}\to\mathcal{A})$ is provable by repeated application of the axioms and the M.P. For this we produce the following sequence of wfs:

1. $\sim\mathcal{B}\to(\sim\mathcal{A}\to\sim\mathcal{B})$ by (A1)

[writing $\sim\mathcal{B}$ for $\mathcal{A}$ and $\sim\mathcal{A}$ for $\mathcal{B}$ in (A1)]

2. $(\sim\mathcal{A}\to\sim\mathcal{B})\to(\mathcal{B}\to\mathcal{A})$ by (A3)

3. $((\sim\mathcal{A}\to\sim\mathcal{B})\to(\mathcal{B}\to\mathcal{A}))\to(\sim\mathcal{B}\to((\sim\mathcal{A}\to\sim\mathcal{B})\to(\mathcal{B}\to\mathcal{A})))$

by (A1)

[writing $(\sim\mathcal{A}\to\sim\mathcal{B})\to(\mathcal{B}\to\mathcal{A})$ for $\mathcal{A}$ and $\sim\mathcal{B}$ for $\mathcal{B}$ in (A1)]

4. $\sim\mathcal{B}\to((\sim\mathcal{A}\to\sim\mathcal{B})\to(\mathcal{B}\to\mathcal{A}))$ by 2, 3 and M.P.

5. $(\sim\mathcal{B}\to((\sim\mathcal{A}\to\sim\mathcal{B})\to(\mathcal{B}\to\mathcal{A})))\to$

$((\sim\mathcal{B}\to(\sim\mathcal{A}\to\sim\mathcal{B}))\to(\sim\mathcal{B}\to(\mathcal{B}\to\mathcal{A})))$ by (A2)

[writing $\sim\mathcal{B}$ for $\mathcal{A}$, $\sim\mathcal{A}\to\sim\mathcal{B}$ for $\mathcal{B}$ and $\mathcal{B}\to\mathcal{A}$ for $\mathcal{C}$ in (A2)]

6. $(\sim\mathcal{B}\to(\sim\mathcal{A}\to\sim\mathcal{B}))\to(\sim\mathcal{B}\to(\mathcal{B}\to\mathcal{A}))$ by 4,5, and M.P.

7. $\sim\mathcal{B}\to(\mathcal{B}\to\mathcal{A})$ by 1, 6 and M.P.

Be convinced that the above sequence of wfs (i.e. 1, 2, 3, 4, 5, 6, 7) is a proof of the Theorem $\sim\mathcal{B}\to(\mathcal{B}\to\mathcal{A})$ of L.

Example 2.2.3.: For any three wfs $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ of L

$$\{\mathcal{A}, \mathcal{B}\to(\mathcal{A}\to\mathcal{C})\} \vdash_{L} \mathcal{B}\to\mathcal{C}$$

This is an example of a **deduction**. We have to show that $\mathcal{B}\to\mathcal{C}$ is deducible from the hypotheses $\mathcal{A}$

and $\mathcal{B} \to (\mathcal{A} \to \mathcal{C})$ by repeated application of the axioms and the M.P.

We have the following sequence:

1. $\mathcal{A}$	hypothesis
2. $\mathcal{B} \to (\mathcal{A} \to \mathcal{C})$	hypothesis
3. $\mathcal{A} \to (\mathcal{B} \to \mathcal{A})$	by (A1)
4. $\mathcal{B} \to \mathcal{A}$	by 1, 3 and M.P.
5. $(\mathcal{B} \to (\mathcal{A} \to \mathcal{C})) \to ((\mathcal{B} \to \mathcal{A}) \to (\mathcal{B} \to \mathcal{C}))$	by (A2)
6. $(\mathcal{B} \to \mathcal{A}) \to (\mathcal{B} \to \mathcal{C})$	by 2, 5 and M.P.
7. $\mathcal{B} \to \mathcal{C}$	by 4, 6 and M.P.

The example illustrates the meaning of a deduction as defined earlier.

For future use as a **standard** result we now give a lemma:

Lemma 2.2.1.: For any wf $\mathcal{A}$ of L

$$\vdash_{L} \mathcal{A} \to \mathcal{A},$$

i.e. for every wf $\mathcal{A}$, $\mathcal{A} \to \mathcal{A}$ is a theorem of L.

Replacing the statement forms occurring in the axioms by appropriate statement forms suitable for our purpose we have the following sequence of wfs:

1. $(\mathcal{A} \to ((\mathcal{A} \to \mathcal{A}) \to \mathcal{A}) \to ((\mathcal{A} \to (\mathcal{A} \to \mathcal{A})) \to (\mathcal{A} \to \mathcal{A}))$ by (A2)
2. $\mathcal{A} \to ((\mathcal{A} \to \mathcal{A}) \to \mathcal{A})$ by (A1)
3. $(\mathcal{A} \to (\mathcal{A} \to \mathcal{A})) \to (\mathcal{A} \to \mathcal{A})$ by 1, 2 and M.P.
4. $\mathcal{A} \to (\mathcal{A} \to \mathcal{A})$ by (A1)
5. $\mathcal{A} \to \mathcal{A}$ by 3, 4 and M.P.

Proposition 2.2.1.(Deduction Theorem): If $\Gamma \cup \{\mathcal{A}\} \vdash_{L} \mathcal{B}$ where $\mathcal{A}$ and $\mathcal{B}$ are wfs of L and Γ is a set of wfs of L(possibly empty), then

$$\Gamma \vdash_{L} \mathcal{A} \to \mathcal{B}.$$

In particular if $\mathcal{A} \vdash_{L} \mathcal{B}$, then $\vdash_{L} \mathcal{A} \to \mathcal{B}$.

Proof: Suppose the deduction

$$\Gamma \cup \{\mathcal{A}\} \vdash_{L} \mathcal{B}$$

consists of a sequence of n wfs

$$\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_n = \mathcal{B} \qquad \ldots\ldots(1)$$

We use induction on the positive integer n. First we prove that the statement is true for n = 1, i.e., for n = 1, $\Gamma \vdash_{L} \mathcal{A} \to \mathcal{B}$.

When n = 1 the sequence (1) contains only one term, namely $\mathcal{B}_1$, and we have three possible cases:

$\mathcal{B}_1$ is an axiom of L

or $\mathcal{B}_1$ is in Γ

or $\mathcal{B}_1$ is $\mathcal{A}$.

Case I : Suppose $\mathcal{B}_1$ is an axiom of L. We propose to show that

$$\Gamma \vdash_{L} \mathcal{A} \to \mathcal{B}_1.$$

For this we produce the following sequence:

1. $\mathcal{B}_1$	axiom of L.
2. $\mathcal{B}_1 \to (\mathcal{A} \to \mathcal{B}_1)$	by(A1)
3. $\mathcal{A} \to \mathcal{B}_1$	by 1, 2 and M.P.

The above sequence means that

$$\vdash_{L} \mathcal{A} \to \mathcal{B}_1$$

i.e., $$\phi \vdash_{L} \mathcal{A} \to \mathcal{B}_1$$

and therefore

$$\Gamma \vdash_{L} \mathcal{A} \to \mathcal{B}_1$$

because $\phi \subset \Gamma$ (See property (1) given at the end of § 2.1). Moreover it is to be noted that in this case $\mathcal{B}_1 = \mathcal{B}$ and so we can write

$$\Gamma \vdash_{L} \mathcal{A} \to \mathcal{B}$$

Case II : Suppose $\mathcal{B}_1$ is in Γ. Here also we have to show that

$$\Gamma \vdash \mathcal{A} \to \mathcal{B}_1, \text{ i.e., } \quad \Gamma \vdash_{L} \mathcal{A} \to \mathcal{B}$$

We have the following sequence:

1. $\mathcal{B}_1$	hypothesis.
2. $\mathcal{B}_1 \to (\mathcal{A} \to \mathcal{B}_1)$	by(A1)

3. $\mathcal{A} \to \mathcal{B}_1$ by 1, 2 and M.P.

Therefore $\mathcal{B} \vdash_L \mathcal{A} \to \mathcal{B}_1$, and so

$$\Gamma \vdash_L \mathcal{A} \to \mathcal{B}_1$$

since $\mathcal{B} \in \Gamma$.

Case III : Suppose $\mathcal{B}_1 = \mathcal{A}$. Here again we have to show that

$$\Gamma \vdash_L \mathcal{A} \to \mathcal{B}_1$$

i.e., $$\Gamma \vdash_L \mathcal{A} \to \mathcal{A}$$

By Lemma 2.1.1. we know that $\Gamma \vdash_L \mathcal{A} \to \mathcal{A}$ and therefore

$$\Gamma \vdash_L \mathcal{A} \to \mathcal{B}_1$$

Thus we have proved that the statement of the proposition is true for n = 1. In order to use induction we now suppose

$$\Gamma \vdash_L \mathcal{A} \to \mathcal{B}_k$$

for $k < i \leq n$, and then prove that

$$\Gamma \vdash_L \mathcal{A} \to \mathcal{B}_i.$$

We have four possibilities to examine :

$\mathcal{B}_i$ is an axiom of L

or $\mathcal{B}_i$ is in Γ

or $\mathcal{B}_i$ is $\mathcal{A}$

or $\mathcal{B}_i$ is obtained from some $\mathcal{B}_j$ and $\mathcal{B}_m$ by using the M.P. where $j < m < i$ and $\mathcal{B}_m$ has the form $\mathcal{B}_j \to \mathcal{B}_i$ (Note that only then you can use the M.P.)

The first three cases are identical to Case I, Case II and Case III respectively considered earlier. So in these cases

$$\Gamma \vdash_L \mathcal{A} \to \mathcal{B}_i.$$

For the last case we have the sequence :

1. $\mathcal{A} \to \mathcal{B}_j$ ($j < i$, and it follows by our supposition)
2. $\mathcal{A} \to \mathcal{B}_m$ ($m < i$, and it follows by our supposition)

3. $\mathcal{A} \to (\mathcal{B}_j \to \mathcal{B}_i)$ (writing $\mathcal{B}_j \to \mathcal{B}_i$ for $\mathcal{B}_m$)

4. $(\mathcal{A} \to (\mathcal{B}_j \to \mathcal{B}_i)) \to ((\mathcal{A} \to \mathcal{B}_j) \to (\mathcal{A} \to \mathcal{B}_i))$ by (A2)

5. $(\mathcal{A} \to \mathcal{B}_j) \to (\mathcal{A} \to \mathcal{B}_i)$ by 3, 4 and M.P.

6. $\mathcal{A} \to \mathcal{B}_i$ by 1, 5 and M.P.

This shows that $\Gamma \vdash_L \mathcal{A} \to \mathcal{B}_i$.

Thus the induction is complete and the case i = n is the required result.

The particular case follows by putting $\Gamma = \phi$.

Proposition 2.2.2.(Converse of the Deduction Theorem) : Let $\mathcal{A}$, $\mathcal{B}$ be two wfs and Γ a set of wfs of L. If $\Gamma \vdash_L \mathcal{A} \to \mathcal{B}$, then

$$\Gamma \cup \{\mathcal{A}\} \vdash_L \mathcal{B}.$$

Proof: Suppose $\Gamma \vdash_L \mathcal{A} \to \mathcal{B}$ and take $\Gamma \cup \{\mathcal{A}\}$ as the set of hypotheses.

Then we have

1. $\mathcal{A} \to \mathcal{B}$	deduction from Γ
2. $\mathcal{A}$	hypothesis in $\Gamma \cup \{\mathcal{A}\}$
3. $\mathcal{B}$	2,1 and M.P.

Therefore $\Gamma \cup \{\mathcal{A}\} \vdash_L \mathcal{B}$.

Proposition 2.2.3. (Hypothetical Syllogism, abbreviated H.S.) : For any three wfs $\mathcal{A}, \mathcal{B}, \mathcal{C}$ of L

$$\{\mathcal{A} \to \mathcal{B}, \mathcal{B} \to \mathcal{C}\} \vdash_L \mathcal{A} \to \mathcal{C}$$

Proof: We take $\{\mathcal{A} \to \mathcal{B}, \mathcal{B} \to \mathcal{C}, \mathcal{A}\}$ as the set of hypotheses and use Deduction Theorem. We have

1. $\mathcal{A} \to \mathcal{B}$	hypothesis
2. $\mathcal{B} \to \mathcal{C}$	hypothesis
3. $\mathcal{A}$	hypothesis
4. $\mathcal{B}$	3, 1 and M.P.
5. $\mathcal{C}$	4, 2 and M.P.

Thus $\{\mathcal{A} \to \mathcal{B}, \mathcal{B} \to \mathcal{C}, \mathcal{A}\} \vdash_L \mathcal{C}$.

Then by Deduction Theorem

$$\{\mathcal{A} \to \mathcal{B}, \mathcal{B} \to \mathcal{C}\} \underset{L}{\vdash} \mathcal{A} \to \mathcal{C}$$

Example 2.2.4. $\{\mathcal{A} \to (\mathcal{B} \to \mathcal{C}), \mathcal{B}\} \underset{L}{\vdash} \mathcal{A} \to \mathcal{C}$

We prove this by Deduction Theorem.

First we show that

$$\{\mathcal{A} \to (\mathcal{B} \to \mathcal{C}), \mathcal{B}, \mathcal{A}\} \underset{L}{\vdash} \mathcal{C}$$

We have the following sequence of wfs :

1. $\mathcal{A}$	hypothesis
2. $\mathcal{B}$	hypothesis
3. $\mathcal{A} \to (\mathcal{B} \to \mathcal{C})$	hypothesis
4. $\mathcal{B} \to \mathcal{C}$	1, 3 and M.P.
5. $\mathcal{C}$	2, 4 and M.P.

Therefore $\{\mathcal{A} \to (\mathcal{B} \to \mathcal{C}), \mathcal{B}, \mathcal{A}\} \underset{L}{\vdash} \mathcal{C}$.

By Deduction Theorem

$$\{\mathcal{A} \to (\mathcal{B} \to \mathcal{C}), \mathcal{B}\} \underset{L}{\vdash} \mathcal{A} \to \mathcal{C}.$$

Example 2.2.5.: For any two wfs $\mathcal{A}$ and $\mathcal{B}$ of L the following is a theorem of L :

$$\sim \mathcal{B} \to (\mathcal{B} \to \mathcal{A})$$

This was proved in an alternative way in Ex.2.2.2. Here we prove it by using the Hypothetical Syllogism (H.S.) :

1. $\sim \mathcal{B} \to (\sim \mathcal{A} \to \sim \mathcal{B})$	(A1)
2. $(\sim \mathcal{A} \to \sim \mathcal{B}) \to (\mathcal{B} \to \mathcal{A})$	(A3)
3. $\sim \mathcal{B} \to (\mathcal{B} \to \mathcal{A})$	H.S.

Lemma 2.2.2 : For any wfs $\mathcal{A}$ and $\mathcal{B}$ of L

$$\{\sim \mathcal{A}\} \underset{L}{\vdash} \mathcal{A} \to \mathcal{B}.$$

The following sequence of wfs proves this :

1. $\sim \mathcal{A}$	hypothesis
2. $\sim \mathcal{A} \to (\sim \mathcal{B} \to \sim \mathcal{A})$	(A1)
3. $\sim \mathcal{B} \to \sim \mathcal{A}$	1, 2 and M.P.
4. $(\sim \mathcal{B} \to \sim \mathcal{A}) \to (\mathcal{A} \to \mathcal{B})$	(A3)

5. $\mathcal{A} \to \mathcal{B}$ 3, 4 and M.P.

Proposition 2.2.4. : For any wf $\mathcal{A}$ of L

$$\vdash_{L} (\sim\mathcal{A} \to \mathcal{A}) \to \mathcal{A}.$$

i.e. $(\sim\mathcal{A} \to \mathcal{A}) \to \mathcal{A}$ is a theorem of L.

Proof : We prove that

$$\{\sim\mathcal{A} \to \mathcal{A}\} \vdash_{L} \mathcal{A}.$$

and then use the Deduction Theorem.

We have the following sequence of wfs:

1. $\sim\mathcal{A} \to \mathcal{A}$	hypothesis
2. $\sim\mathcal{A} \to (\sim\sim(\sim\mathcal{A} \to \mathcal{A}) \to \sim\mathcal{A})$	(A1)
3. $(\sim\sim(\sim\mathcal{A} \to \mathcal{A}) \to \sim\mathcal{A}) \to (\mathcal{A} \to \sim(\sim\mathcal{A} \to \mathcal{A}))$	(A3)
4. $\sim\mathcal{A} \to (\mathcal{A} \to \sim(\sim\mathcal{A} \to \mathcal{A}))$	2, 3 and H.S.
5. $(\sim\mathcal{A} \to (\mathcal{A} \to \sim(\sim\mathcal{A} \to \mathcal{A}))) \to$ $((\sim\mathcal{A} \to \mathcal{A}) \to (\sim\mathcal{A} \to \sim(\sim\mathcal{A} \to \mathcal{A})))$	(A2)
6. $(\sim\mathcal{A} \to \mathcal{A}) \to (\sim\mathcal{A} \to \sim(\sim\mathcal{A} \to \mathcal{A}))$	4, 5 and M.P.
7. $\sim\mathcal{A} \to \sim(\sim\mathcal{A} \to \sim\mathcal{A})$	1, 6 and M.P.
8. $(\sim\mathcal{A} \to \sim(\sim\mathcal{A} \to \mathcal{A})) \to ((\sim\mathcal{A} \to \mathcal{A}) \to \mathcal{A})$	(A3)
9. $(\sim\mathcal{A} \to \mathcal{A}) \to \mathcal{A}$	7, 8 and M.P.
10. $\mathcal{A}$	1, 9 and M.P.

Hence $\{\sim\mathcal{A} \to \mathcal{A}\} \vdash_{L} \mathcal{A}$ which gives on using Deduction Theorem :

$$\vdash_{L} (\sim\mathcal{A} \to \mathcal{A}) \to \mathcal{A}$$

Example 2.2.6. : For any wf $\mathcal{A}$ of L

$$\vdash_{L} \sim\sim\mathcal{A} \to \mathcal{A}$$

i.e. $\sim\sim\mathcal{A} \to \mathcal{A}$ is a theorem of L.

To eastablish this it is sufficient to show that

$$\{\sim\sim\mathcal{A}\} \vdash_{L} \mathcal{A}$$

because the use of Deduction Theorem will then give the required result.

We have

1. $\sim\sim\mathcal{A}$	hypothesis
2. $\sim\mathcal{A} \rightarrow \mathcal{A}$	taking $A = \sim\mathcal{A}$ and $\mathcal{B} = \mathcal{A}$ in Lemma 2.2.2.
3. $(\sim\mathcal{A} \rightarrow \mathcal{A}) \rightarrow \mathcal{A}$	Proposition 2.2.4.
4. $\mathcal{A}$	2, 3 and M.P.

Example 2.2.7 : For any three wfs $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ of L

$$\text{(i)} \vdash_L \mathcal{A} \rightarrow ((\mathcal{A} \rightarrow \mathcal{B}) \rightarrow \mathcal{B})$$

$$\text{(ii)} \vdash_L (\mathcal{A} \rightarrow (\mathcal{B} \rightarrow \mathcal{C})) \rightarrow (\mathcal{B} \rightarrow (\mathcal{A} \rightarrow \mathcal{C}))$$

(i) We prove that $\{\mathcal{A}, \mathcal{A} \rightarrow \mathcal{B}\} \vdash_L \mathcal{B}$ and use Deduction Theorem **twice**.

1. $\mathcal{A}$	hypothesis
2. $\mathcal{A} \rightarrow \mathcal{B}$	hypothesis
3. $\mathcal{B}$	1, 2 and M.P.

Therefore $\{\mathcal{A}, \mathcal{A} \rightarrow \mathcal{B}\} \vdash_L \mathcal{B}$.

By deduction Theorem

$$\{\mathcal{A}\} \vdash_L (\mathcal{A} \rightarrow \mathcal{B}) \rightarrow \mathcal{B}$$

Using Deduction Theorem once again

$$\vdash_L \mathcal{A} \rightarrow ((\mathcal{A} \rightarrow \mathcal{B}) \rightarrow \mathcal{B})$$

(ii) We prove that $\{\mathcal{A}, \mathcal{B}, \mathcal{A} \rightarrow (\mathcal{B} \rightarrow \mathcal{C})\} \vdash_L \mathcal{C}$

and use Deduction Theorem **thrice**.

1. $\mathcal{A}$	hypothesis
2. $\mathcal{B}$	hypothesis
3. $\mathcal{A} \rightarrow (\mathcal{B} \rightarrow \mathcal{C})$	hypothesis
4. $\mathcal{B} \rightarrow \mathcal{C}$	1, 3 and M.P.
5. $\mathcal{C}$	2, 4 and M.P.

Therefore $\{\mathcal{A}, \mathcal{B}, \mathcal{A} \rightarrow (\mathcal{B} \rightarrow \mathcal{C})\} \vdash_L \mathcal{C}$

Then using the Deduction theorem **thrice** :

$$\{\mathcal{B}, \mathcal{A} \rightarrow (\mathcal{B} \rightarrow \mathcal{C})\} \vdash_L \mathcal{A} \rightarrow \mathcal{C}$$

$$\{\mathcal{A} \to (\mathcal{B} \to \mathcal{C})\} \vdash_{L} \mathcal{B} \to (\mathcal{A} \to \mathcal{C})$$

$$\vdash_{L} (\mathcal{A} \to (\mathcal{B} \to \mathcal{C})) \to (\mathcal{B} \to (\mathcal{A} \to \mathcal{C}))$$

Example 2.2.8. : For three wfs $\mathcal{A}, \mathcal{B}, \mathcal{C}$ of L

$$\{\mathcal{A}\} \vdash_{L} (\mathcal{A} \to (\mathcal{B} \to \mathcal{C})) \to (\mathcal{B} \to \mathcal{C})$$

As in Ex 2.2.7. prove that

$$\{\mathcal{A}, \mathcal{B}, \mathcal{A} \to (\mathcal{B} \to \mathcal{C})\} \vdash_{L} \mathcal{C}$$

and use Deduction Theorem **twice.**

Summary :

- A proof is a sequence of wfs $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, ..., \mathcal{A}_n$ such that each $\mathcal{A}_i$ is an axiom or is deducible from some of the preceding wfs and the M.P.
- A theorem $\mathcal{A}$ is a wf such that there is a proof

  $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, ..., \mathcal{A}_n$ with $\mathcal{A} = \mathcal{A}_n$ (notation $\vdash_{L} \mathcal{A}$)
- A deduction $\mathcal{A}$ from a set Γ of wfs is a sequence $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, ..., \mathcal{A}_n$ of wfs such that $\mathcal{A} = \mathcal{A}_n$ where each $\mathcal{A}_i$ is

  (i) an axiom **or** (ii) a wf in Γ **or** (iii) deducible from some of the preceding wfs and the M.P.
- **Deduction Theorem :**

  If $\Gamma \cup \{\mathcal{A}\} \vdash_{L} \mathcal{B}$ where $\mathcal{A}$ and $\mathcal{B}$ are wfs of L and Γ is a set of wfs of L(possibly empty)

  then $\Gamma \vdash_{L} \mathcal{A} \to \mathcal{B}$.
- **Converse of the Deduction Theorem :**

  Let $\mathcal{A}, \mathcal{B}$ be two wfs and Γ a set of wfs of L. If $\Gamma \vdash_{L} \mathcal{A} \to \mathcal{B}$, then $\Gamma \cup \{\mathcal{A}\} \vdash_{L} \mathcal{B}$.
- **Hypothetical Syllogism :**

  For any three wfs $\mathcal{A}, \mathcal{B}, \mathcal{C}$ of L

  $$\{\mathcal{A} \to \mathcal{B}, \mathcal{B} \to \mathcal{C}\} \vdash_{L} \mathcal{A} \to \mathcal{C}$$

Unit-3

The Adequacy Theorem for L

Introduction : Unit 1 gave us a notion of 'logical truth', namely that of tautology. It would be reasonable to hope that these logical truths will correspond to the theorems of L, and to attempt to construct L with this end in view.

3.1 The Adequacy Theorem for L : Though the symbols of the language of L are being thought of purely as formal symbols, L was defined in such a way that we could interpret the wfs. of L as statement forms and that then each truth function is represented by some wf. Thus although we cannot talk of assigning truth values to the symbols of L i precisely the same way as in Unit 1, we can define an analogous procedure.

Definition : A *valuation* of L is a function v whose domain is the set of wfs, of L and whose range is the set $\{T, F\}$ such that, for any wfs. $\mathcal{A}, \mathcal{B}$ of L.

(i) $v(\mathcal{A}) \neq v(\sim\mathcal{A})$, and

(ii) $v(\mathcal{A} \to \mathcal{B}) = F$ if and only if $v(\mathcal{A}) = T$ and $v(\mathcal{B}) = F$

Note that an arbitrary 'assignment of truth values' to the symbols $p_1, p_2, ...$ of L will yield a valuation, as each *wf.* of L will (as a statement form) take one of the two truth values under such an assignment. (i) and (ii) will then obviously be satisfied.

Definition : A *wf.* $\mathcal{A}$ of L is a tautology if for every valuation v, $v(\mathcal{A}) = T$. This is same as regarding $\mathcal{A}$ as a statement form and applying the previous definition.

Proposition 3.1.1. (The Soundness Theorem) : Every theorem of L is a tautology (in the truth-table sense)

Proof: To prove this proposition it is sufficient to verify that

(1) all the axioms of L are tautologies,

(2) if $\mathcal{A}$ and $\mathcal{A} \to \mathcal{B}$ are tautologies, then so is $\mathcal{B}$,

because a theorem is provable by repeated application of the axioms and the Modus Ponens.

The result (2) was proved in a proposition in UNIT 1. The axioms of L are:

$$\mathcal{A} \to (\mathcal{B} \to \mathcal{A})$$

$$(\mathcal{A} \to (\mathcal{B} \to \mathcal{C})) \to ((\mathcal{A} \to \mathcal{B}) \to (\mathcal{A} \to \mathcal{C}))$$

$$(\sim\mathcal{A} \to \sim\mathcal{B}) \to (\mathcal{B} \to \mathcal{A})$$

By constructing truth-tables one can easily verify that these are tautologies.

We now give some definitions and propositions with the primary aim to define a **tautology** in the formal theory L and establish another famous theorem known as the **Adequacy Theorem** of Propositional Calculus. It will appear that a tautology in the truth-table sense as defined in UNIT 1 is also a tautology in this new sense.

An **extension** L^* of L is a formal theory obtained by altering or enlarging the set of axioms of L so that all theorems of L are also theorems of L^* (which may have some new theorems).

An extension L^* of L is **consistent** if for no wf $\mathcal{A}$ of L are both $\mathcal{A}$ and $\sim\mathcal{A}$ theorems of L^*.

Note that L can be considered an extension of itself, and it can be proved that L is consistent as an extension of itself.

Proposition 3.1.2. : L is consistent as an extension of itself.

Proof : Suppose L is not consistent. Then there exists a wf $\mathcal{A}$ of L such that both $\mathcal{A}$ and $\sim\mathcal{A}$ are theorems of L. By Soundness Theorem then both $\mathcal{A}$ and $\sim\mathcal{A}$ are tautologies. But this is impossible, because if $\mathcal{A}$ is a tautology, then $\sim\mathcal{A}$ is a contradiction.

Proposition 3.1.3. : An extension L^* of L is consistent if and only if there is a wf of L which is not a theorem of L^*.

Proof : Let L^* be consistent. Then for any wf $\mathcal{A}$ both $\mathcal{A}$ and $\sim\mathcal{A}$ cannot be theorems i.e. either $\mathcal{A}$ is not a theorm or $\sim\mathcal{A}$ is not a theorem of L^* so that L^* has at least one wf which is not a theorem.

To prove the converse it is sufficient to show that if L^* is not consistent, then every wf is a theorem of L^*.

If L^* is not consistent then $\vdash_{L^*}\mathcal{B}$ and $\vdash_{L^*}\sim\mathcal{B}$ for some wf $\mathcal{B}$. Let $\mathcal{A}$ be any wf.

First we prove that

$$\vdash_{L}\sim\mathcal{B}\rightarrow(\mathcal{B}\rightarrow\mathcal{A})$$

For this, it is sufficient to prove that

$$\{\sim\mathcal{B}\}\vdash_{L}\mathcal{B}\rightarrow\mathcal{A}$$

and then apply the Deduction Theorem :

1. $\sim\mathcal{B}$	hypothesis
2. $\sim\mathcal{B}\rightarrow(\sim\mathcal{A}\rightarrow\sim\mathcal{B})$	(A1)
3. $\sim\mathcal{A}\rightarrow\sim\mathcal{B}$	1, 2 and M.P.
4. $(\sim\mathcal{A}\rightarrow\sim\mathcal{B})\rightarrow(\mathcal{B}\rightarrow\mathcal{A})$	(A3)
5. $\mathcal{B}\rightarrow\mathcal{A}$	3, 4 and M.P.

By Deduction theorem we get

$$\vdash_{L}\sim\mathcal{B}\rightarrow(\mathcal{B}\rightarrow\mathcal{A})$$

But L^* is an extension of L. So we have

$$\vdash_{L^*} \sim\mathscr{B}\rightarrow(\mathscr{B}\rightarrow\mathscr{A})$$

Therefore in L^* we have

1. $\mathscr{B}$	Theorem
2. $\sim\mathscr{B}$	Theorem
3. $\sim\mathscr{B}\rightarrow(\mathscr{B}\rightarrow\mathscr{A})$	Theorem
4. $\mathscr{B}\rightarrow\mathscr{A}$	2, 3, and M.P.
5. $\mathscr{A}$	1, 4 and M.P.

Thus, $\vdash_{L^*}\mathscr{A}$, i.e., every wf $\mathscr{A}$ is a Theorem in L^*

This completes the proof.

Proposition 3.1.4. : Let L^* be a consistent extension of L and let $\mathscr{A}$ be a wf of L which is not a theorem of L^*. Then L^{**} is also consistent where L^{**} is the extension of L obtained from L^* by including $\sim\mathscr{A}$ as an additional axiom.

Proof : Given that $\mathscr{A}$ is a wf of L which is not a theorem of L^*. Suppose L^{**} is not consistent. Then for some $\mathscr{B}$

$$\vdash_{L^{**}}\mathscr{B}$$

and $\vdash_{L^{**}}\sim\mathscr{B}$

We can also prove as in Proposition 2.2.7 that

$$\vdash_{L^{**}}\mathscr{A}$$

But L^{**} differs from L^* only in having $\sim\mathscr{A}$ as an additional axiom. So $\vdash_{L^{**}}\mathscr{A}$ is equivalent to

$$\{\sim\mathscr{A}\}\vdash_{L^*}\mathscr{A}$$

By Deduction Theorem

$$\vdash_{L^*}\sim\mathscr{A}\rightarrow\mathscr{A} \qquad (1)$$

By Proposition 2.2.4.

$$\vdash_{L}(\sim\mathscr{A}\rightarrow\mathscr{A})\rightarrow\mathscr{A}$$

and therefore

$$\vdash_{L^*} (\sim\mathcal{A} \rightarrow \mathcal{A}) \rightarrow \mathcal{A} \qquad (2)$$

By (1), (2) and M.P. we get

$$\vdash_{L^*} \mathcal{A}$$

But this contradicts the hypothesis that $\mathcal{A}$ is not a theorem of L^*. Hence L^{**} must be consistent.

An extension L^* of L is said to be **complete** if for each wf $\mathcal{A}$ either $\mathcal{A}$ or $\sim\mathcal{A}$ is a theorem of L^*.

Proposition 3.1.5. : Let L^* be a consistent extension of L. Then there is a consistent complete extension of L^*.

Proof : See A.G.Hamilton's Logic for Mathematicians, pages 41, 42.

Since L is a consistent extension of itself, **L has a consistent complete extension.**

A **valuation** v of L is a function from the set of all wfs of L to the set {T, F} such that

(i) $v(\mathcal{A}) \neq v(\sim\mathcal{A})$

and (ii) $v(\mathcal{A} \rightarrow \mathcal{B}) = F$ if and only if $v(\mathcal{A}) = T$ and $v(\mathcal{B}) = F$.

Note that if we consider a wf as a statement form and {T, F} as the set of truth-values TRUE and FALSE, then a truth function dicussed earlier is a valuation.

A wf $\mathcal{A}$ of L is a **tautology** if for every valuation v, $v(\mathcal{A}) = T$.

It is very important to appreciate that a tautology in the truth-table sense is a tautology in this sense also. This follows immediately from the above **formal** definition of tautology, and so we have the following Proposition in view of Proposition 3.1.1.

Proposition 3.1.6. : Every theorem of L is a tautology (in the formal sense).

Proposition 3.1.7. : If L^* is a consistent extension of L then there is a valution in which each theorem of L^* takes the value T.

Proof : Let L_c^{**} be a consistent complete extension of L^*(Such an L_c^{**} exists in view of Proposition 3.1.6.)

Define v on the set of wfs of L by

$$v(\mathcal{A}) = T \text{ if } \vdash_{L_c^{**}} \mathcal{A}$$

and $$v(\mathcal{A}) = F \text{ if } \vdash_{L_c^{**}} \sim\mathcal{A}$$

for any wf $\mathcal{A}$ [since L_c^{**} is complete either $\vdash_{L_c^{**}} \mathcal{A}$ or $\vdash_{L_c^{**}} \sim\mathcal{A}$, and so we can define v as above].

We now want to show that v is a valuation, i.e.,

(i) $v(\mathcal{A}) \neq v(\sim \mathcal{A})$

(ii) $v(\mathcal{A} \to \mathcal{B}) = F$ if and only if $v(\mathcal{A}) = T$ and $v(\mathcal{B}) = F$.

(i) follows from the fact that L_c^{**} is consistent, i.e., there is no $\mathcal{A}$ for which $\mathcal{A}$ and $\sim \mathcal{A}$ are both theorems of L_c^{**}. For (ii) we proceed as follows :

Suppose first $v(\mathcal{A}) = T$, $v(\mathcal{B}) = F$ and suppose on the contrary $v(\mathcal{A} \to \mathcal{B}) = T$, then

$\vdash_{L_c^{**}} \mathcal{A}$ (definition of v) (1)

$\vdash_{L_c^{**}} \sim \mathcal{B}$ (assumption and definition of v) (2)

$\vdash_{L_c^{**}} \mathcal{A} \to \mathcal{B}$ (assumption and definition of v) (3)

By (1), (3) and M.P.

$\vdash_{L_c^{**}} \mathcal{B}$ (4)

From (2) and (4) we get a contradiction to the consistency of L_c^{**} ($\mathcal{B}, \sim \mathcal{B}$ cannot both be theorems). Hence our assumption $v(\mathcal{A} \to \mathcal{B}) = T$ is not correct, and so, if $v(\mathcal{A}) = T$ and $v(\mathcal{B}) = F$, then $v(\mathcal{A} \to \mathcal{B}) = F$.

Conversely suppose $v(\mathcal{A} \to \mathcal{B}) = F$ and suppose on the contrary either $v(\mathcal{A}) = F$ or $v(\mathcal{B}) = T$. Then

$\vdash_{L_c^{**}} \sim(\mathcal{A} \to \mathcal{B})$ (assumption and definition of v)

and either

$\vdash_{L_c^{**}} \sim \mathcal{A}$ (assumption and definition of v) (5)

or $\vdash_{L_c^{**}} \mathcal{B}$ (assumption and definition of v) (6)

Now

$\vdash_{L_c^{**}} \sim \mathcal{A} \to (\sim \mathcal{B} \to \sim \mathcal{A})$ (A1) (7)

and $\vdash_{L_c^{**}} \mathcal{B} \to (\mathcal{A} \to \mathcal{B})$ (A1) (8)

By (5), (7) and M.P. or (6), (8) and M.P. we have

$\vdash_{L_c^{**}} \sim \mathcal{B} \to \sim \mathcal{A}$ (9)

or $\vdash_{L_c^{**}} \mathcal{A} \to \mathcal{B}$

Now

$$\vdash_{L_c^{**}} (\sim \mathcal{B} \to \sim \mathcal{A}) \to (\mathcal{A} \to \mathcal{B}) \quad \text{(A3)} \qquad (10)$$

But by (9), (10) and M.P. we also have

$$\vdash_{L_c^{**}} \mathcal{A} \to \mathcal{B}$$

Therefore in either case

$$\vdash_{L_c^{**}} \mathcal{A} \to \mathcal{B}$$

But we already saw that

$$\vdash_{L_c^{**}} \sim(\mathcal{A} \to \mathcal{B})$$

This contradicts the consistency of L_c^{**}. Thus $v(\mathcal{A} \to \mathcal{B}) = F$ must imply $v(\mathcal{A}) = T$ and $v(\mathcal{B}) = F$. Therefore v is a valuation.

Now let $\mathcal{A}$ be a theorem of L^*. Then $\vdash_{L_c^{**}} \mathcal{A}$ because L_c^{**} is an extension of L^*.

Therefore $v(\mathcal{A}) = T$.

Proposition 3.1.8. (Adequacy or Completeness Theorem) : If a wf $\mathcal{A}$ of L is a tautology (in the formal sense) then it is a theorem of L.

Proof : Let $\mathcal{A}$ be a wf of L such that $\mathcal{A}$ is a tautology. Suppose $\mathcal{A}$ is not a theorem. Let L^* be the extension of L obtained by including the additional axiom $\sim\mathcal{A}$. Then by a slight modification of Proposition 2.2.8. (taking L for L^*) L^* is consistent. Therefore by Proposition 2.2.11.there is a valuation v which gives every theorem of L^* the value T. In particular

$v(\sim\mathcal{A}) = T$ (Note that $\sim\mathcal{A}$ is an axiom, and so as good as a theorem)

But $v(\mathcal{A}) = T$ because $\mathcal{A}$ is a tautology. Thus we get a contradiction to the fact that v is a valuation $[v(\sim\mathcal{A}) \neq v(\mathcal{A})]$. Hence $\mathcal{A}$ is a theorem of L.

Summary :

- **The Soundness Theorem :** Every theorem of *L* is a tautology.
- L is consistent as an extension of itself.
- An extension L^* of L is consistent if and only if there is a wf of L which is not a theorem of L^*
- If a wf $\mathcal{A}$ of L is a tautology (in the formal sense), then it is a theorem of L.

G.U. Questions

1997

1. Define a **Theorem** of formal propositional calculus L. Prove that for any wf $\mathcal{A}$, $\mathcal{A} \to \mathcal{A}$ is a theorem of L. $1 + 4 = 5$

2. Prove that in the formal theory L, $\mathcal{A} \to \mathcal{C}$ is deducible from $\mathcal{A} \to \mathcal{B}$ and $\mathcal{B} \to \mathcal{C}$ where $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ are any wfs. 5

3. Explain clearly the meaning of a tautology in the formal theory L. Prove that if a wf $\mathcal{A}$ of L is a tautology, then it is a theorem of L. $2 + 9 = 11$

1998

1. Give a set of axioms for formal theory L of propositional calculus.

   Prove that the following statement form is a theorm of L:

   $(\mathcal{A} \to \mathcal{B}) \to (\mathcal{A} \to \mathcal{A})$ $2 + 3 = 5$

2. When is a theory L^* said to be an **extension** of L? When do you call L^* consistent?

   Prove that an extension L^* of L is consistent if and only if there is a well-formed formula which is not a theorem in L^*. $1 + 1 + 6 = 8$

3. Explain how a tautology **in the truth-table sense** may be considered as a tautology in the formal theory L. 3

4. Prove that the statement form $\mathcal{A} \to \mathcal{C}$ is a consequence in L of the set of statement forms $\{\mathcal{A} \to \mathcal{B}, \mathcal{B} \to \mathcal{A}\}$ 5

1999

1. If $\Gamma \cup \{\mathcal{A}\} \vdash_{L} \mathcal{B}$ where $\mathcal{A}$ and $\mathcal{B}$ are well-formed formulas(wfs) of an axiomatic theory L of propositional calculus and Γ is a set of wfs of L(possibly empty) then prove that

   $$\Gamma \vdash_{L} \mathcal{A} \to \mathcal{B}$$

   Deduce that if $\mathcal{A} \vdash_{L} \mathcal{B}$, then $\vdash_{L} \mathcal{A} \to \mathcal{B}$. 8

2. Define an **extension** L^* of an axiomatic theory L of propositional calculus. When is an extension L^* of L said to be consistent?

   Prove that L, as its own extension, is consistent. $1 + 1 + 3 = 5$

3. Define a **tautology** in an axiomatic theory L of propositional calculus using the notion of valuations of L.

   If L^* is a consistent extension of L then prove that there is a valuaton in which each theorem of L^* takes the value T. Hence deduce that a tautology in L is a theorem of L. $2 + 7 + 2 = 11$

Unit 4
Informal Predicate Calculus

Introduction : In Unit 1 we analysed sentences and arguments, breaking them down into constituent simple statements, regarding these simple statements as the building blocks. By this means we were able to discover something of what makes a valid argument. The symbol of $\mathcal{E}$ may be interpreted in many different ways, but we shall now concern ourselves with the purely formal aspects of the language and consider the logical relationships of wfs, rather than properties which depend on particular interpretations.

4.1. Symbolism of Predicate Calculus :

Let us consider the following sentences :

1. Hilton is a human being.
2. Mina is a friend of Rita.
3. A bird has wings.
4. The successor of an integer is an integer.
5. The product of two numbers is a number.

In predicate calculus such sentences are symbolised as follows :

1. H(h), h — Hilton, H — is a human being.
2. F(m, r), m — Mina, r — Rita, F — is a friend of
3. W(b), b — a bird, W — has wings.
4. I(s(x)), x — any integer, s(x) — successor of an integer x, I - is an integer
5. R(p(y, z)), y, z — any two numbers, p(y, z) — the product of the numbers y and z, R — is a number.

The letters used in the above symbolism fall into three categories in regard to their meanings :

(A) H, F, W I, R — each of them denotes a **certain property**.

(B) h, m, r — each of them denotes a **fixed** object.

b, x, y, z — each of them denotes **any object of a certain class,** i.e., a **variable** object.

(C) s, p — each of them denotes a **function** in the mathematical sense.

The **capital letters** denoting some properties are called **predicate letters.**

The **small letters** denoting fixed objects are called **individual constants.**

The small letters denoting variables objects are called **individual variables.**

The small letters denoting functions are called **function letters.**

Example 4.1.1. :

Symbolise the following using predicate letters, individual variables and constants and function letters

(a) A bird and an animal have legs.

(b) The rational numbers are real numbers.

(c) 2 divides 4.

(d) A rose or a dahlia is pink.

(e) The square of an even number is even.

(f) The square of a number is divisible by the number.

(a) The sentence is equivalent to :

A bird has legs and an animal has legs.

b—a bird, a—an animal, L—has legs.

$L(b) \wedge L(a)$

L—predicate letter

a, b—individual variables.

(b) The sentence is equivalent to :

If x is a rational number, then x is a real number.

Q—is a rational number.

R—is a real number

$(Q(x) \rightarrow R(x))$

Q, R—predicate letters

x—individual variable

(c) D(2, 4) D—divides

D—predicate letter

2, 4—individual constants

(d) r—rose, d—dahlia, P—is pink

$P(r) \vee P(d)$

P—predicate letter

r, d—individual variables

(e) E—is even, e—even number

s(e)—square of e

E(s(e)), E—predicate letter, e—individual variable,

s—function letter

(f) n—a number

s(n)—the square of n

D—is divisible by

D(s(n), n)

D—predicate letter

n—individual variable

s—function letter

The quantifiers $\forall$ and $\exists$:

Consider the statement—

Every rational number is a real number.

The statement is equivalent to—

For all x, if x is a rational number, then x is a real number.

This may be expressed in predicate calculus symbolism as follows—

For all x, $(Q(x) \rightarrow R(x))$ ......(1)

where Q stands for "is a rational number" and R for "is a real number".

Consider again the statement—

All men are mortal, which is equivalent to—

For all x, if x is a man, then x is mortal. In predicate calculus symbolism—

For all x, $(A(x) \rightarrow M(x))$ ....(2)

where A means "is a man" and M means "is mortal".

The phrase **for all x** is called a universal **quantifier** and symbolised as $(\forall x)$.

The statements (1) and (2) above can now be written as

$$(\forall x)(Q(x) \rightarrow R(x))$$

and $(\forall x)(A(x) \rightarrow M(x))$ respectively.

Now consider the statement—

Some real numbers are rational, which is equivalent to—

There exists at least one real number such that it is rational, i.e., **There exists at least one object x such that** x is a real number and x is a rational number.

Similarly the statement—

Some pigs have wings may be written as

There exists at least one object x such that x is a pig and x has wings.

The phrase "there exists at least one object x" is called an **existential quantifier** symbolised as $(\exists x)$.

The above two statements may now be symbolised as

$$(\exists x)(R(x) \wedge Q(x))$$

$$(\exists x)(P(x) \wedge W(x))$$

with R(x), Q(x), P(x) and W(x) having obvious meanings.

Example 4.1.2. :

Translate the following sentences into symbols using quantifiers, predicate letters, individual variables etc.

(1) All judges are lawyers.

(2) Not all lawyers are judges.

(3) Judge Sinha is neither active nor honest.

(4) Some writers admire women.

(5) No woman is both a politician and a housewife.

Solution :

(1) For all x, if x is a judge, then x is a lawyer.

J—is a judge

L—is a lawyer

$(\forall x)(J(x) \rightarrow L(x))$

(2) ~ (All lawyers are judges)

$\sim (\forall x)(L(x) \rightarrow J(x))$

(3) s—Judge Sinha, A—is active, H—is honest.

$\sim A(s) \wedge \sim H(s)$

Judge Sinha is not active and Judge Sinha is not honest.

(4) There exists an x such that x is a writer and x admires women.

$(\exists x)(W(x) \wedge A(x))$

(5) ~ (For all x, x is a politician and x is a housewife).

$\sim (\forall x)(P(x) \wedge H(x))$

Example 4.1.3. :

Translate into symbols :

(a) Not all birds can fly.

(b) Anyone can do that.

(c) Some people are stupid.

Solution :

(a) $\sim (\forall x)(B(x) \rightarrow F(x))$

(b) $(\forall x)(M(x) \rightarrow C(x))$

(c) $(\exists x)(M(x) \wedge S(x))$

Let us look at (a) more closely—

It is not true that for all x if x is a bird then x can fly.

The sentence is equivalent to—

There are some birds which cannot fly. i.e., there is an x such that x is a bird and x cannot fly:

$$(\exists x)(B(x) \wedge \sim F(x)).$$

This shows that the existential quantifier $(\exists x)$ can be replaced by $\sim (\forall x)$.

In general the following two sentences have the same meaning:

(i) There exists an x which has property P.

(ii) It is not true that all x's do not have property P.

Thus

$$(\exists x)P(x) \text{ and } \sim (\forall x)(\sim P(x))$$

are equivalent in meaning.

Also from our knowledge of propositional calculus we know that for any two statements $\mathcal{A}$ and $\mathcal{B}$

$(\mathcal{A} \wedge \mathcal{B})$ is equivalent to $(\sim(\mathcal{A} \rightarrow (\sim\mathcal{B})))$,

$(\mathcal{A} \vee \mathcal{B})$ is equivalent to $((\sim\mathcal{A}) \rightarrow \mathcal{B})$,

$(\mathcal{A} \leftrightarrow \mathcal{B})$ is equivalent to $(\mathcal{A} \rightarrow \mathcal{B}) \wedge (\mathcal{B} \rightarrow \mathcal{A})$,

in the sense that each pair has the same truth-value. So from our discussion we can dispense with the symbols

$$\exists, \wedge, \vee \text{ and } \leftrightarrow.$$

4.2. First order language

The symbolic expressions of the type discussed above constitute what is called **First Order language** $\mathcal{L}$.

If we have to deal with sequences of individual variables and constants, predicate letters and function letters, then it is convenient to use some conventions in the symbolism.

Let A be a predicate letter such that we have a symbol of the form $A(s_1, s_2, \ldots, s_n)$. Then A is called a predicate letter of n **arguments** and is usually denoted by A^n. If we have to deal with sequences of predicate letters of different arguments, then we use symbols like:

$A_1^1, A_2^1, A_3^1, \ldots$	(predicate letters with 1 argument)
$A_1^2, A_2^2, A_3^2, \ldots$	(predicate letters with 2 arguments)
....	
$A_1^n, A_2^n, A_3^n, \ldots$	(predicate letters with n arguments)
....	

Again let f be a function letter such that we have a symbol of the form $f(t_1, t_2, \ldots, t_n)$. Then f is called a function letter of n **arguments** and is usually denoted by f^n. While dealing with of such function letters of different arguments we use the symbols:

$f_1^1, f_2^1, \ldots$	(function letters with 1 argument)
$f_1^2, f_2^2, \ldots$	(function letters with 2 arguments)

....

$f_1^n, f_2^n, \ldots$ (function letters with n arguments)

Sequences of individual variables and constants are denoted by

$x_1, x_2, \ldots\ldots$

and $a_1, a_2, \ldots\ldots$

respectively.

The following set of symbols is called the **alphabet** of the First Order Language :

$x_1, x_2, \ldots\ldots$	individual variables
$a_1, a_2, \ldots\ldots$	individual constants
$A_1^1, A_2^1, \ldots, A_1^2, A_2^2, A_3^2, \ldots$	predicate letters
$f_1^1, f_2^1, \ldots, f_1^2, f_2^2, \ldots$	function letters
(,), ,	punctuation symbols
$\sim, \rightarrow$	connectives
$\forall$	quantifier

Example 4.2.1. :

Let x_1, x_2 be natural numbers. Express '$x_1 + x_2 = x_1x_2$' in first order language.

We use predicate letter A_1^2 for =, function letters f_1^2 and f_2^2 for + and × respectively. Then

$f_1^2(x_1, x_2)$	stands for $x_1 + x_2$
$f_2^2(x_1, x_2)$	stands for x_1x_2
$A_1^2(f_1^2(x_1, x_2), f_2^2(x_1, x_2))$	stands for $x_1 + x_2 = x_1x_2$.

Example 4.2.2. :

Let x_1 be an element of a group.

Express $x_1x_1^{-1}$ = identity, in first order language.

Let a_1 stand for the identity

f_1^1 stand for the function which takes each element to its inverse

f_1^2 stand for the product of two group elements

A_1^2 stand for equality between two elements.

Then

$f_1^1(x_1)$	stands for x_1^{-1}.
$f_1^2(x_1, f_1^1(x_1))$	stands for $x_1x_1^{-1}$
$A_1^2(f_1^2(x_1, f_1^1(x_1)), a_1)$	stands for $x_1x_1^{-1}$ = identity.

Definition :

Let $\mathcal{L}$ be a first order language.

A **term** in $\mathcal{L}$ is defined as follows :

(i) Individual constants and variables are terms.

(ii) If f_i^n is a function letter in $\mathcal{E}$, and $t_1, t_2, \ldots, t_n$ are terms in $\mathcal{E}$, then $f_i^n(t_1, t_2, \ldots, t_n)$ is a term in $\mathcal{E}$.

(iii) The set of all terms is generated as in (i) and (ii).

In **Example 4.2.2.** above

$a_1, x_1, f_1^1(x_1)$ and $f_1^2(x_1, f_1^1(x_1))$ are terms.

Definition :

An **atomic formula** in $\mathcal{E}$ is an expression of the form $A_j^k(t_1, t_2, \ldots, t_k)$ where A_j^k is a predicate letter and $t_1, t_2, \ldots, t_k$ are terms in $\mathcal{E}$.

In **Example 4.2.2.** $A_1^2(f_1^2(x_1, f_1^1(x_1)), a_1)$ is an atomic formula of predicate calculus.

Definition :

A **well-formed formula** (wf) of $\mathcal{E}$ is defined as follows :

(i) Every atomic formula of $\mathcal{E}$ is a wf of $\mathcal{E}$

(ii) If $\mathcal{A}$ and $\mathcal{B}$ are wfs of $\mathcal{E}$ by (i) so are

$(\sim \mathcal{A}), (\mathcal{A} \rightarrow \mathcal{B})$ and $(\forall x)\mathcal{A}$, where x is any variable

(iii) The set of all wfs of $\mathcal{E}$ is generated as in (i) and (ii).

Example 4.2.3. :

The following are atomic formulas :

(1) $A_1^2(f_1^2(x_1, x_2), f_2^2(x_1, x_2))$ (Type (i))

(2) $\sim (A_1^1(x_1) \rightarrow A_2^1(x_1))$ (Type (ii))

$(\forall x_1)(A_1^1(x_1) \rightarrow A_2^1(x_1))$

(3) $(\forall x_1)(\forall x_2)(A_1^2(x_1, x_2) \rightarrow A_1^1(x_2))$ (Type (iii))

$(\forall x_1)A_1^2(x_1, x_2) \rightarrow (\forall x_3)A_2^2(x_3, x_1)$

Definition :

In the wf $(\forall x_i)\mathcal{A}$, we say that $\mathcal{A}$ is the **scope** of the quantifier $(\forall x_i)$. Moreover generally, when $((\forall x_i)\mathcal{A})$ occurs as a **subformula** of a wf $\mathcal{B}$ we say that the scope of $(\forall x_i)$ in $\mathcal{B}$ is $\mathcal{A}$.

Example 4.2.4. :

In the wf $(\forall x_1)A_1^1(x_2)$, the scope of the quantifier $(\forall x_1)$ is $A_1^1(x_2)$.

Example 4.2.5. :

In the wf $(\forall x_1)(\forall x_2)(A_1^2(x_1, x_2)) \rightarrow A_1^1(x_2))$ the scope of the quantifier $(\forall x_2)$ is

$(A_1^2(x_1, x_2) \rightarrow A_1^1(x_2))$ and the scope of the quantifier $(\forall x_1)$ is $(\forall x_2)(A_1^2(x_1, x_2) \rightarrow A_1^1(x_2))$.

Example 4.2.6. :

The scope of $(\forall x_2)$ in $(\forall x_1)(A_1^2(x_1, x_2) \rightarrow (\forall x_2) A_1^1(x_2))$ is $A_1^1(x_2)$.

Definition :

We say that a variable x_i **occurs bound if it** occurs in the scope of a $(\forall x_i)$ or in $(\forall x_i)$ of a wf. If a variable x_i does not occur bound then it is said to **occur free.**

Example 4.2.7. :

In the wf $(\forall x_1) A_1^1(x_2)$, the variable x_1 occurs bound but the variable x_2 occurs free.

Example 4.2.8. :

In the wf $(\forall x_1)(\forall x_2)(A_1^2(x_1, x_2) \rightarrow A_1^1(x_2))$ both the variables x_1 and x_2 occur bound.

Example 4.2.9. :

In the wf $(\forall x_1)(A_1^2(x_1, x_2) \rightarrow (\forall x_2) A_1^1(x_2))$ the variable x_1 occurs bound twice and x_2 occurs free once and bound twice.

Definition :

Let $\mathscr{A}$ be any wf of $\mathscr{L}$. A term t is said to be **free for x_i in $\mathscr{A}$** if x_i **does not occur free** in $\mathscr{A}$ within the scope of a $(\forall x_j)$ where x_j is some variable occurring in t.

Example 4.2.10. :

Let $\mathscr{A}$ be the wf

$$((\forall x_1) A_1^2(x_1, x_2) \rightarrow (\forall x_3) A_2^2(x_3, x_1))$$

in $\mathscr{L}$. Consider the following terms

$$f_1^2(x_1, x_4), f_2^2(x_2, x_3), x_2, f_4^2(x_1, x_3).$$

(1) $f_1^2(x_1, x_4)$ is **not free for x_2** in $\mathscr{A}$, because x_2 occurs free within the scope of $(\forall x_1)$.

(2) $f_2^2(x_2, x_3)$ is free for x_3 in $\mathscr{A}$, because x_3 does not occur free in $\mathscr{A}$ within the scope of $(\forall x_3)$.

(3) x_2 is free for x_1 in $\mathscr{A}$, because x_1 does not occur free in $\mathscr{A}$ within the scope of $(\forall x_2)$ [in fact there is no $(\forall x_2)$ in $\mathscr{A}$]

(4) $f_4^2(x_1, x_3)$ is not free for x_1 in $\mathscr{A}$, because x_1 occurs free in the scope of $(\forall x_3)$.

Example 4.2.11. :

Which occurences of x_1 in the following wfs are free and which are bound?

(a) $(\forall x_2)(A_1^2(x_1, x_2) \rightarrow A_1^2(x_2, a_1))$

(b) $(A_1^1(x_3) \rightarrow (\sim (\forall x_1)(\forall x_2) A_1^3(x_1, x_2, a_1)))$

(c) $((\forall x_1) A_1^1(x_1) \rightarrow (\forall x_2) A_1^2(x_1, x_2))$

(d) $(\forall x_2)(A_1^2(f_1^2(x_1, x_2), x_1) \rightarrow (\forall x_1) A_2^2(x_3, f_2^2(x_1, x_2))$

Is the term $f_1^2(x_1, x_2)$ free for x_2 in any or all of these?

First Part :

(a) The only occurence of x_1 is free.

(b) Both the occurences of x_1 are bound (not free)

(c) The first two occurences of x_1 are bound and the third occurence is free.

(d) The first two occurences of x_1 are free and the third and the fourth occurences are bound.

Second part :

(a) The variables in the given term are x_1 and x_2.

The given wf contains a $(\forall x_2)$ in the scope of which x_2 occurs bound, i.e. does not occur free.

Therefore $f_1^2(x_1, x_2)$ is free for x_2 in the given wf.

(b) The given wf contains a $(\forall x_2)$ in the scope of which x_2 occurs bound, i.e. does not occur free.

Therefore $f_1^2(x_1, x_2)$ is free for x_2 in the given wf.

(c) Argument is same as (b).

(d) Argument is same as (b).

4.3. Interpretation :

Let us **recall** the following example—

Let x_1, x_2 be natural numbers. The equality

$$x_1 + x_2 = x_1 x_2$$

can be expressed in first order language as follows :

$$A_1^2(f_1^2(x_1, x_2), f_2^2(x_1, x_2))$$

when we assign the following intrepretations to the symbols used in our first order language :

the variables x_1, x_2	stand for natural numbers
the function letter f_1^2	stands for the **function** +
the function letter f_2^2	stands for the **function** ×
the predicate letter A_1^2	stands for the **relation** =

Again we saw that the relation $x_1 x_1^{-1}$ = identity in a group may be expressed in first order language as

$$A_1^2(f_1^2(x_1, f_1^1(x_1)), a_1)$$

if we give the following interpretations to the symbols used :

the variable x_1	stands for any element of a group
the constant a_1 of a group	stands for a distinguished element
the function letter f_1^2	stands for the function "product"
the function letter f_2^2	stands for the function "inverse"
the predicate letter A_1^2	stands for the relation =

In general a wf of $\mathcal{E}$ contains some of the following symbols :

(1) $x_1, x_2, \ldots\ldots$	individual variables
(2) $a_1, a_2, \ldots\ldots$	individual constants
(3) $f_1^1, f_2^1, \ldots, f_1^2, f_2^2, \ldots$	function letters
(4) $A_1^1, A_2^1, \ldots, A_1^2, A_2^2, \ldots$	predicate letters
(5) (,), ,	punctution symbols
(6) $\sim, \rightarrow$	connectives
(7) $\forall$	quantifiers

The meaning or the interpretation of the symbols at (5), (6) and (7) is known to us in any situation. But in order to know the meaning or interpretation of a wf of $\mathcal{E}$ we must assign proper meaning or interpretation to the symbols at (1) — (4). Roughly speaking an **interpretation** of wf of $\mathcal{E}$ consists of a "set of meanings" of

(1) the variables

(2) the constants

(3) the function letters

and (4) the predicate letters that occur in the wf.

Definition :

An **interpretation** I of $\mathcal{E}$ consists of

(1) a non-empty set D_I called the **domain** of I

(2) a collection of **distinguished** elements

$$\{\bar{a}_1, \bar{a}_2, \ldots\} \text{ of } D_I$$

(3) a collection of **functions** $\{\bar{f}_i^{\,n}, i > 0, n > 0\}$ on D_I

and (4) a collection of **relations** $\{\bar{A}_i^{\,n}, i > 0, n > 0\}$ on D_I

Example 4.3.1. :

Consider the wf

$$(\forall x_1)(\forall x_2)(\sim (\forall x_3)(\sim A_1^2(f_1^2(x_1, x_3), x_2)))$$

Translate the above wf into everyday language with the following interpretation, and state the truth or falsity of the statement representing the wf :

(1) the domain $D_I = \{0, 1, 2, 3, ...\}$

(2) $\bar{f}_1^2$ is +

(3) $\bar{A}_1^2$ is =

With this interpretation the given wf is the statement :

"for all whole numbers x and y it is not true that for any whole number z, $x+z \neq y$";

alternatively,

"for any two whole numbers x and y there exists a whole number z such that

$x+z=y$"

We know that this statement is **false**.

Example 4.3.2. :

Consider again the same wf as in Ex.3.3.1. :

$$(\forall x_1)(\forall x_2)(\sim(\forall x_3)(\sim A_1^2(f_1^2(x_1, x_3), x_2)))$$

Translate the above wf into everyday language with the following interpretation, and state the truth or falsity of the statement representing the wf :

(1) D_I is the set of all positive rational numbers;

(2) $\bar{f}_1^2$ is multiplication, $\bar{f}_2^2$ is division

(3) $\bar{A}_1^2$ is equality.

with this interpretation the given wf becomes the statement :

"For any two positive rational numbers x and y there is a positive rational number z, such that

$xz=y$".

This is a known property of the rational numbers.

The above two examples reveal a very important property of wfs in first order language :

The truth or falsity of a statement representing a wf in first order language depends on the interpretation.

Example 4.3.3. :

Let $\mathcal{L}$ be a first order language which includes (besides variables, punctuation symbols, connectives and quantifier) the individual constant a_1, the function letter f_1^2 and the predicate letter A_1^2. Let $\mathcal{B}$ denote the wf

$$(\forall x_1)(\forall x_2)(A_1^2(f_1^2(x_1, x_2), a_2) \to A_1^2(x_1, x_2)).$$

Define an interpretation I of $\mathcal{L}$ as follows :

(1) D_I is Z, the set of integers;

(2) $\bar{a}_1$ is 0;

(3) $\bar{f}_1^2$ is - ;

(4) $\bar{A}_1^2$ is <.

Write down the interpretation (meaning in everyday language) of $\mathcal{A}$ in I. Is this a true statement or a false one? Find another interpretation in which $\mathcal{A}$ is interpretated by a statement with the opposite truth value.

With the given interpretation the wf means :

For any two integers x and y, if x - y is less than 0, then x is less than y.

This is **true**.

Now consider the following interpretation :

(1) D_I is Z;

(2) $\bar{a}_1$ is 0;

(3) $\bar{f}_1^2$ is + ;

(3) $\bar{A}_1^2$ is =.

With this interpretation the given wf means :

For any two integers x and y, if x + y is equal to 0, then x is equal to y.

This is **false**.

Example 4.3.4. :

Is there an interpretation in which the wf

$$(\forall x_1)(A_1^1(x_1) \rightarrow A_1^1(f_1^1(x_1)))$$

of $\mathcal{E}$ can be interpreted by a false statement?

Consider the following interpretation I :

(1) D_I is Z;

(2) $\bar{a}_1$ is 0;

(3) $\bar{f}_1^1$ is the negative of an element of D_I;

(4) $\bar{A}_1^1$ is the relation "is greater than 0".

With the above interpretation the given wf means—

For any integer x, if x is greater than 0 then - x is also greater than zero.

This is **false**.

Note : In all the above examples, the variables in each wf occur all bound. In each case with a given interpretation the statement representating the wf is either **true** or **false**. If a wf contains free variables then we face an uncertainty regarding truth or falsity of the statement as will be seen in the examples below :

Example 4.3.5. :

Consider the following wf of $\mathcal{B}$:

$$A_1^2(f_1^2(x_1, x_2), f_2^2(x_1, x_2))$$

Here both the variables occur free.

Define an interpretation I as follows :

(1) $D_I = \{1, 2, 3, ...\}$

(2) $\bar{f}_1^2$ is addition, $\bar{f}_2^2$ is multiplication

(3) $\bar{A}_1^2$ is equality

Then the given wf becomes the statement

$$x + y = xy$$

where x, y are positive integers.

The above statement is not true for **any two** elements x, y of the domain **but it is true** for the pair (2, 2) of the elements of the domain. Such a pair (2, 2) will be called a **satisfaction** of the wf under the above interpretation.

Example 4.3.6. :

Consider the wf of $\mathcal{B}$:

$$(\forall x_1)(A_1^2(f_1^2(x_1, x_2), a_1))$$

The wf contains a free variable x_2.

Define an interpretation I as follows :

(1) D_I is Z

(2) $\bar{a}_1$ is 0

(3) $\bar{f}_1^2$ is multiplication

(4) $\bar{A}_1^2$ is equality

Under the above interpretation the given wf becomes the statement—

For every integer x, xy = 0, where y is an integer. A **satisfaction** for the given wf under the above interpretation is (m, 0) where m is any integer.

4.4. Satisfiability and truth :

Definition :

Let I be an interpretation with the domain D_I and let $\mathcal{T}$ be the set of all terms of $\mathcal{B}$.

A **valuation** v in I is a function

$$v : \mathcal{T} \rightarrow D_I$$

with the properties—

(i) $v(t_i) = \bar{t}_i$ for each term t_i of $\mathcal{E}$.

(ii) $v(f_i^n(t_1, t_2, ..., t_n)) = \bar{f}_i^n(v(t_1), ..., v(t_n))$ where $t_1, ..., t_n$ are any terms of $\mathcal{E}$ and f_i^n is any function letter of $\mathcal{E}$.

A valuation is thus a rule which assigns to each term in $\mathcal{E}$ the object in D_I which is to be its interpretation. Part (ii) above ensures that the rule is a consistent one.

Remarks :

(1) In general, in a given interpretation there may exist **many valuations**.

(2) A given valuation will assign an element of D_I to each of the variables x_i of $\mathcal{E}$. A valuation v will be **completely specified** by giving $v(x_1), v(x_2), ...$ This is because the $v(x_i)$ are given by the definition and, inductivity, for any term $f_i^n(t_1, t_2, ..., t_n)$, the value of $v(f_i^n(t_1, t_2, ..., t_n))$ is determined by (ii).

Definition :

Two valuations v and v′ in the same interpretation are called **i-equivalent** if $v(x_j) = v'(x_j)$ for every $j \neq i$.

Definition :

Let $\mathcal{A}$ ba a wf of $\mathcal{E}$ and let I be an interpretation of $\mathcal{E}$. $\mathcal{A}$ valuation v in I is said to **satisfy** the wf $\mathcal{A}$ if it can be shown inductivily to do so under the following four conditions :

(i) v satisfies the atomic formula $A_j^n(t_1, ..., t_n)$ if $\bar{A}_j^n(v(t_1), ..., v(t_n))$ is true in D_I.

(ii) v sastisfies $\sim \mathcal{B}$ if v does not satisfy $\mathcal{B}$.

(iii) v satisfies $(\mathcal{B} \rightarrow \mathcal{C})$ if either v satisfies $\sim \mathcal{B}$ or v satisfies $\mathcal{C}$.

(iv) v satisfies $(\forall x_i)\,\mathcal{B}$ if every valuation v′ which is i-equivalent to v satisfies $\mathcal{B}$.

Remark 1 :

For any v and $\mathcal{A}$, either v satisfies $\mathcal{A}$ or v satisfies $\sim \mathcal{A}$.

Remark 2. :

Roughly speaking by saying that a valuation v **satisfies** the wf $\mathcal{A}$ of $\mathcal{E}$ we mean the following :

Replace each term t in $\mathcal{A}$ by v(t) and replace each function letter and predicate letter by its interpretation in I. We obtain a statement about the elements of D_I.

This statement is now either **true** or **false**. If it is true, then we say that v satisfies $\mathcal{A}$.

Example 4.4.1. :

Consider the wf of $\mathcal{E}$

$$A_1^2(f_1^2(x_1, x_2), f_1^2(x_3, x_4))$$

under the interpretation I defined as follows :

(1) D_I is $N = \{1, 2, 3, ...\}$

(2) $\bar{f}_1^2$ is multiplication

(3) $\bar{A}_1^2$ is equality

Under the interpretation I the given wf becomes the statement

$$x_1x_2 = x_3x_4 \text{ where } x_1, x_2, x_3, x_4 \in N.$$

Define a valuation v in I such that

$$v(x_1) = 2, v(x_2) = 6, v(x_3) = 3, v(x_4) = 4,$$

Then v satisfies the given wf. But let us take another valuation w in I such that

$$w(x_1) = 1, w(x_2) = 5, w(x_3) = 4, w(x_4) = 2$$

Then $1 \times 5 \neq 4 \times 2$. So w does not satisfy the given wf.

Example 4.4.2. :

Consider the wf

$$(\forall x_1) A_1^2(f_1^2(x_1, x_2), f_1^2(x_2, x_1))$$

of $\mathcal{L}$ under the interpretation I defined as follows :

(1) D_I is $N = \{1, 2, 3, ...\}$

(2) $\bar{f}_1^2$ is multiplication

(3) $\bar{A}_1^2$ is equality

The given wf is interpreted as "for every $n \in D_I$, $nm = mn$", which is obviously true.

Let v be a valuation in I.

Note that

(1) the atomic formula $A_1^2(f_1^2(x_1, x_2), f_1^2(x_2, x_1))$ is then interpreted as

$$v(x_1) \times v(x_2) = v(x_2) \times v(x_1)$$

which is certainly true. So v satisfies $A_1^2(f_1^2(x_1, x_2), f_1^2(x_2, x_1))$.

(2) If v′ is another valuation which is i-equivalent to v, then v′ also satisfies $A_1^2(f_1^2(x_1, x_2), f_1^2(x_2, x_1))$.

So by (iv) of the definition above, v satisfies $(\forall x_1) A_1^2(f_1^2(x_1, x_2), f_1^2(x_2, x_1))$.

Thus every valuation v in I satisfies this wf.

Example 4.4.3.. :

Consider the wf

$$(\forall x_1) A_1^2(x_1, a_1)$$

and the interpretation I of Example 3.4.2., viz

(1) D_I is N;

(2) $\bar{a}_1$ is 0

(3) $\bar{A}_1^2$ is =.

Let us take a valuation v in I. Then $A_1^2(x_1, a_1)$ is interpretated as

$$v(x_1) = 0$$

Let us take a valuation v' which is i-equivalent to v, i.e., $v'(x_1) \neq v(x_1)$. But we see that v' does not satisfy $A_1^2(x_1, a_1)$. Therefore by (iv) of the definition of satisfiability, v does not satisfy $(\forall x_1) A_1^2(x_1, a_1)$. Since v is arbitrary, we see that no valuation satisfies the wf.

Definition :

A wf $\mathcal{A}$ is said to be **true** in an interpretation I if **every** valuation in I satisfies $\mathcal{A}$, and $\mathcal{A}$ is said to be **false if no valuation** in I satisfies $\mathcal{A}$.

The wf of Example 3.4.2 is true in the given interpretation.

The wf of Example 3.4.3. is false.

Notation :

We use $I \models \mathcal{A}$ to mean "$\mathcal{A}$ is true in I".

Note that $\models$ is not a symbol of first order language, but it is a part of the **metalanguage**.

It may happen that for a particular wf $\mathcal{A}$, some valuations in I satisfiy $\mathcal{A}$ and some valuations do not. Such a wf is **neither** true **nor** false in I. See Ex. 3.4.3. above.

It is clear from the definition of valuation that a given valuation either satisfies or does not satisfy a given wf $\mathcal{A}$, and hence it is impossible for a wf to be both true and false in a given interpretation.

In a given interpretation I, a wf $\mathcal{A}$ is false if and only if $\sim\mathcal{A}$ is true. This follows from the condition (ii) of the definition of satisfiability.

Proposition 4.4.1. :

In a given interpretation I, a wf $\mathcal{A} \rightarrow \mathcal{B}$ is false if and only if $\mathcal{A}$ is true and $\mathcal{B}$ is false.

Proof :

First suppose $\mathcal{A} \rightarrow \mathcal{B}$ is false in I.

Then no valuation satisfies $\mathcal{A} \rightarrow \mathcal{B}$ in I.

Given any valuation v, then, v does not satisfy $\mathcal{A} \rightarrow \mathcal{B}$. Condition (iii) in the definition of satisfiability says that v satisfies $\mathcal{A} \rightarrow \mathcal{B}$ if either v satisfies $\sim\mathcal{A}$ or v satisfies $\mathcal{B}$. Therefore v does not satisfy $\mathcal{A} \rightarrow \mathcal{B}$ if v does not satisfy $\sim\mathcal{A}$ and v does not satisfy $\mathcal{B}$ i.e., if v satisfies $\mathcal{A}$ and v does not satisfy $\mathcal{B}$ i.e., if $\mathcal{A}$ is true and $\mathcal{B}$ is false.

Next suppose $\mathcal{A}$ is true and $\mathcal{B}$ is false. Let v be any valuation in I. Then v satisfies $\mathcal{A}$ and v does not satisfy $\mathcal{B}$. Therefore v does not satisfy $\sim\mathcal{A}$ and v does not satisfy $\mathcal{B}$. Hence by condition (iii) of the definition of satisfiability v does not satisfy $\mathcal{A} \rightarrow \mathcal{B}$, i.e., $\mathcal{A} \rightarrow \mathcal{B}$ is false.

Proposition 4.4.2. :

If in a particular interpretation I, the wfs $\mathcal{A}$ and $\mathcal{A} \to \mathcal{B}$ are true then $\mathcal{B}$ is also true.

Proof :

Let v be any valuation in I. Suppose $\mathcal{A}$ and $\mathcal{A} \to \mathcal{B}$ are true. Then

v satisfies $\mathcal{A}$ ......(1)

and v satisfies $\mathcal{A} \to \mathcal{B}$ ......(2)

Then from (2) by applying condition (iii) of the definition of satisfiability, either v satisfies $\sim \mathcal{A}$ or v satisfies $\mathcal{B}$. But by (1) v does not satisfy $\sim \mathcal{A}$. So v satisfies $\mathcal{B}$.

Proposition 4.4.3. :

Let $\mathcal{A}$ be a wf of $\mathcal{L}$, and let I be an interpretation of $\mathcal{L}$. Then $I \models \mathcal{A}$ if and only if $I \models (\forall x_i)\,\mathcal{A}$ where x_i is any variable.

Proof :

Suppose $I \models \mathcal{A}$, i.e., $\mathcal{A}$ is true in I. Then v satisfies $\mathcal{A}$. Let v′ be any valuation which is i-equivalent to v. Then v′ also satisfies $\mathcal{A}$ (since every valuation satisfies $\mathcal{A}$, $\mathcal{A}$ being true). Hence by condition (iv) of the definition of satisfiability v satisfies $(\forall x_i)\,\mathcal{A}$, i.e., $I \models (\forall x_i)\,\mathcal{A}$.

Next suppose $I \models (\forall x_i)\,\mathcal{A}$. Then v satisfies $(\forall x_i)\,\mathcal{A}$. Hence every v′ which is i-equivalent to v satisfies $\mathcal{A}$ (by condition (iv) of the definition of satisfiability). In particular, v satisfies $\mathcal{A}$, and so every valuation satisfies $\mathcal{A}$, i.e. $I \models \mathcal{A}$.

By repeated application of the above theorem we get the following corollary.

Corollary :

Let $y_1, y_2, ..., y_n$ be variables in $\mathcal{L}$, let $\mathcal{A}$ be a wf of $\mathcal{L}$, and let I be an interpretation. Then $I \models \mathcal{A}$ if and only if $I \models (\forall y_1)\,(\forall y_2)...(\forall y_n)\,\mathcal{A}$.

In first order language we use only two connectives $\sim$ and $\to$ and one quantifier $(\forall x)$. The connectives $\wedge$ and $\vee$, and the quantifier $(\exists x)$ are **defined** as symbols as follows :

$(\exists x_i)\mathcal{A}$ is an abbreviation for $(\sim((\forall x_i)(\sim \mathcal{A})))$

$(\mathcal{A} \wedge \mathcal{B})$ is an abbreviation for $(\sim(\mathcal{A} \to (\sim \mathcal{B})))$

$(\mathcal{A} \vee \mathcal{B})$ is an abbreviation for $((\sim \mathcal{A}) \to \mathcal{B})$

Proposition 4.4.4. :

In an interpretation I, a valuation v satisfies the formula $(\exists x_i)\mathcal{A}$ if and only if there is at least one valuation v′ which is i-equivalent to v and which satisfies $\mathcal{A}$.

Proof :

$(\exists x_i)\mathcal{A}$ stands for $(\sim((\forall x_i)\sim(\mathcal{A})))$.

Let v satisfy $(\sim((\forall x_i)(\sim\mathcal{A})))$. Then v does not satisfy $(\forall x_i)(\sim\mathcal{A})$. By condition (iv) of the definition of satisfiability there is some valuation v′ which is i-equivalent to v and which does not satisfy $(\sim\mathcal{A})$. This v′ then must satisfy $\mathcal{A}$.

Conversely suppose v′ satisfies $\mathcal{A}$. Then v′ does not satisfy $\sim\mathcal{A}$. By condition (iv) of the definition of satisfiability v does not satisfy $(\forall x_i)(\sim\mathcal{A})$, and therefore v satisfies $\sim((\forall x_i)(\sim\mathcal{A}))$.

Definition :

Let $\mathcal{A}_0$ be a wf of L (formal theory of propositional calculus). Replace each statement letter of $\mathcal{A}_0$ by a wf of $\mathcal{L}$ (first order language) replacing the same statement letter by the same wf of $\mathcal{L}$ throughout.

Then we get a wf $\mathcal{A}$ of $\mathcal{L}$. This wf $\mathcal{A}$ of $\mathcal{L}$ is called a **substitution instance** of $\mathcal{A}_0$ in $\mathcal{L}$.

Example 4.4.4. :

Consider the wf

$$\mathcal{A}_0 = (\sim p_1 \rightarrow (p_2 \rightarrow p_3)) \text{ of L}$$

Substitute

$(\forall x_1)A_1^1(x_1)$	for	p_1
$(\forall x_2)A_1^2(x_1, x_2)$	for	p_2
$(A_1^1(x_1) \rightarrow (\forall x_1)A_2^1(x_1))$	for	p_3

Then

$$\mathcal{A} = \sim(\forall x_1)A_1^1(x_1) \rightarrow ((\forall x_2)A_1^2(x_1, x_2)) \rightarrow (A_1^1(x_2) \rightarrow (\forall x_1)A_2^1(x_1))$$

is a substitution instance of $\mathcal{A}_0$ in L.

Note : $\mathcal{A}$ is also a substitution instance of $\mathcal{A}_0^1 = \sim p_1 \rightarrow p_2$ because $\mathcal{A}$ is obtained by substituting $(\forall x_1)A_1^1(x_1)$ for p_1, and $(\forall x_2)A_1^2(x_1, x_2) \rightarrow (A_1^1(x_2) \rightarrow (\forall x_1)A_2^1(x_1))$ for p_2.

Definition :

$\mathcal{A}$ wf of $\mathcal{L}$ is called a **tautology** (in $\mathcal{L}$) if it is a substistution instance in $\mathcal{L}$ of a tautology in L.

It can be proved (see Hamilton, p 65) that **a wf $\mathcal{A}$ of $\mathcal{L}$ which is a tautology is true in any interpretation of $\mathcal{L}$.**

It can also be proved that if $\mathcal{A}$ is a wf of $\mathcal{L}$ in which there is no free variable (such a wf is said to be **closed**), and I is any interpretation of $\mathcal{A}$ then

either $I \models \mathcal{A}$ or $I \models \sim\mathcal{A}$ i.e. $\mathcal{A}$ is either true or false in I. This fact was already illustrated by examples.

Definition :

A wf $\mathcal{A}$ of $\mathcal{L}$ is said to be **logically valid** if $\mathcal{A}$ is true in every interpretation of $\mathcal{L}$. A wf $\mathcal{A}$ of $\mathcal{L}$ is said to be **contradictory** if it is false in every interpretation.

Proposition 4.4.5. :

If the wfs $\mathcal{A}$ and $\mathcal{A} \rightarrow \mathcal{B}$ of $\mathcal{L}$ are logically valid, then $\mathcal{B}$ is logically valid.

Proof :

First prove that (Proposition 3.4.2.) if in a particular interpretation the wfs $\mathcal{A}$ and $\mathcal{A} \rightarrow \mathcal{B}$ are true, then $\mathcal{B}$ is also true. Then apply the definition of logical validity.

Proposition 4.4.6. :

$\mathcal{A}$ wf $\mathcal{A}$ of $\mathcal{L}$ is logically valid if and only if $(\forall x_i)\,\mathcal{A}$ is logically valid, where x_i is any variable.

Proof :

Prove first that (Proposition 3.4.3) in any interpretation in $\mathcal{L}$, $\mathcal{A}$ is true if and only if $(\forall x_i)\,\mathcal{A}$ is true for any variable x_i. Then apply the definition of logical validity.

Remark :

1. To show that a wf $\mathcal{A}$ is logically valid, we just have to prove that an **arbitrary** valuation in an **arbitrary** interpretation satisfies $\mathcal{A}$.

2. To show that a wf $\mathcal{A}$ is not logically valid, we must construct an interpretation in which there is a valuation which does not satisfy it.

Definition :

A wf $\mathcal{A}$ of $\mathcal{L}$ said to **logically imply** another wf $\mathcal{B}$ of $\mathcal{L}$ if a valuation v satisfies $\mathcal{A}$, then v satisfies $\mathcal{B}$ also in every interpretation.

$\mathcal{A}$ wf $\mathcal{B}$ of $\mathcal{L}$ is said to be a logical consequence of a set Γ of wfs of $\mathcal{L}$ if and only if in every interpretation every valuation that satisfies every wfΓ also satisfies $\mathcal{B}$.

Two wfs $\mathcal{A}$ and $\mathcal{B}$ of $\mathcal{L}$ are said to be **logically equivalent** if and only if they logically imply each other.

Definition :

An interpretation I is said to be a **model** for a set Γ of wfs of $\mathcal{L}$ if and only if every wf of Γ is true for I.

Example 4.4.5. :

(1) The wf

$$(\forall x_1)(\exists x_2)A_1^2(x_1, x_2) \to (\exists x_1)(\forall x_2)A_1^2(x_1, x_2)$$

is not logically valid.

We have to choose an interpretation I, and a valuation v which does not satisfy the wf.

Let I be an interpretation such that

(1) D_I is Z

(2) $\overline{A}_1^2(x_1, x_2)$ is $x_1 < x_2$.

Then $(\forall x_1)(\exists x_2)A_1^2(x_1, x_2)$ is true but $(\exists x_1)(\forall x_2)A_1^2(x_1, x_2)$ is false for any valuation.

(2) Let $\mathcal{A}$ be a closed wf of $\mathcal{L}$. Then we know that in any interpretation I of $\mathcal{L}$ either $\mathcal{A}$ is true or $\sim\mathcal{A}$ is true. Therefore the wf

$$\mathcal{A} \vee (\sim\mathcal{A})$$

is always true in any interpretation I of $\mathcal{L}$.

Formal Predicate Calculus :

4.5. A Set of AXIOMS for Predicate Calculus :

Let $\mathcal{L}$ be a first order language. **Formal** or **axiomatic predicate calculus** is a deductive system denoted $K_{\mathcal{L}}$ or simply K, with the following axiom schemas and rules of interference(deduction).

Axiom schemas :

Let $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ be any wfs of $\mathcal{L}$. The following are axiom schemas of $K_{\mathcal{L}}$:

(K1) $(\mathcal{A} \to (\mathcal{B} \to \mathcal{A}))$

(K2) $(\mathcal{A} \to (\mathcal{B} \to \mathcal{C})) \to ((\mathcal{A} \to \mathcal{B}) \to (\mathcal{A} \to \mathcal{C}))$

(K3) $(\sim\mathcal{A} \to \sim\mathcal{B}) \to (\mathcal{B} \to \mathcal{A})$

(K4) $((\forall x_i)\ \mathcal{A} \to \mathcal{A})$, if x_i does not occur free in $\mathcal{A}$

(K5) $((\forall x_i)\ \mathcal{A}(x_i) \to \mathcal{A}(t))$, if $\mathcal{A}(x_i)$ is a wf of $\mathcal{L}$ and t is a term in $\mathcal{L}$ which is free for x_i in $\mathcal{A}(x_i)$.

(K6) $(\forall x_i)(\mathcal{A} \to \mathcal{B}) \to (\mathcal{A} \to (\forall x_i)\ \mathcal{B})$, if $\mathcal{A}$ contains no free occurence of the variable x_i.

Rules of Inference (Deduction) :

1. Modus Ponens(MP) :

From $\mathcal{A}$ and $(\mathcal{A} \to \mathcal{B})$, deduce $\mathcal{B}$.

2. Generalisation :

From $(\forall x_i)\ \mathcal{A}$ deduce $\mathcal{A}$ for any variable x_i.

Definition :

(1) A **proof** in $K_{\mathcal{L}}$ is a finite sequence of wfs $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n$ of $\mathcal{L}$ such that for each $i(1 \leq i \leq n)$, either $\mathcal{A}_i$ is an axiom of $K_{\mathcal{L}}$ or $\mathcal{A}_i$ is deduced from the previous members of the sequence by MP or Generalisation.

(2) If Γ is a set of wfs of $\mathcal{L}$, a **deduction from** Γ in $K_{\mathcal{L}}$ is a similar sequence in which members of Γ may be included.

(3) A wf $\mathcal{A}$ is a **theorem** of $K_{\mathcal{L}}$ if it is the last member of some sequence which constitues a proof in $K_{\mathcal{L}}$.

(4) A wf $\mathcal{A}$ is a **consequence in** $K_{\mathcal{L}}$ of the set Γ of wfs if $\mathcal{A}$ is the last member of a sequence which constitues a deduction from Γ in $K_{\mathcal{L}}$.

Notation :

(1) $\underset{K_{\mathcal{L}}}{\vdash} \mathcal{A}$ means "$\mathcal{A}$ is a theorem of $K_{\mathcal{L}}$."

(2) $\Gamma \underset{K_{\mathcal{L}}}{\vdash} \mathcal{A}$ means "$\mathcal{A}$ is a consequence of Γ in $K_{\mathcal{L}}$" where Γ is a set of wfs of $K_{\mathcal{L}}$.

(3) Very often (when there is no scope for ambiguity) K means $K_{\mathcal{L}}$

Proposition 4.5.1. :

If $\mathcal{A}$ is a wf of $\mathcal{L}$ and $\mathcal{A}$ is a tautology, then $\mathcal{A}$ is a theorem of K.

Proof :

Recall that a wf $\mathcal{A}$ of $\mathcal{L}$ is a tautology if there is a wf $\mathcal{A}_1$ of L from which $\mathcal{A}$ is obtained by substituting wfs of $\mathcal{L}$ for the statement letters, and which is a tautology. Let $\mathcal{A}$ be a wf of $\mathcal{L}$ which is a tautology, and let $\mathcal{A}_0$ be the corresponding wf of L. Then $\mathcal{A}_0$ is a tautology of L and so $\underset{L}{\vdash} \mathcal{A}_0$. The proof of $\mathcal{A}_0$ in L can be transformed into a proof of $\mathcal{A}$ in K simply by replacing statement letters by appropriate wfs of $\mathcal{L}$ throughout, because the axiom schems (L1), (L2), (L3) and the MP are common to both the systems L and K.

Note :

In constrast to the situation in propositionnal calculus the converse of the above theorem is not true.

Proposition 4.5.2. :

All instances of (K1), (K2) and (K3) are logically valid.

Proof :

We already mentioned that a tautology in $\mathcal{L}$ is true in any interpretation i.e., a tautology in $\mathcal{L}$ is logically valid. (K1), (K2) and (K3) are tautologies because their propositional calculus counterparts (L1), (L2) and (L3) are tautologies. Therefore (K1), (K2) and (K3) are logically valid.

Proposition 4.5.3. :

All instances of the axiom schemas (K4), (K5) and (K6) are logically valid, i.e., if $\mathcal{A}$ and $\mathcal{B}$ are any wfs of $\mathcal{L}$, then the following are logically valid :

(a) $((\forall x_i)\, \mathcal{A} \rightarrow \mathcal{A})$ if x_i does not occur free in $\mathcal{A}$.

(b) $((\forall x_i)\, \mathcal{A}(x_i) \rightarrow \mathcal{A}(t))$, if $\mathcal{A}(x_i)$ is a wf of $\mathcal{L}$ and t is a term in $\mathcal{L}$ which is free for x_i in $\mathcal{A}(x_i)$.

(c) $((\forall x_i)(\mathcal{A} \rightarrow \mathcal{B}) \rightarrow (\mathcal{A} \rightarrow (\forall x_i)\, \mathcal{B}))$, if $\mathcal{A}$ contains no free occurence of the variable x_i.

Proof :

(a) Let I be any interpretation of $\mathcal{L}$ and v any valuation in I.

We know (from Proposition 3.4.1.) that

$((\forall x_i)\, \mathcal{A} \rightarrow \mathcal{A})$ is false if and only if

$(\forall x_i)\, \mathcal{A}$ is true and $\mathcal{A}$ is false.

Suppose $(\forall x_i)\, \mathcal{A}$ is true, so that v satisfies $(\forall x_i)\, \mathcal{A}$. Then by condition (iv) of the definition of satisfiability, every v′ which is i-equivalent to v satisfies $\mathcal{A}$. In particular v also satisfies $\mathcal{A}$. This means that v is true in I. Therefore $((\forall x_i)\, \mathcal{A} \rightarrow \mathcal{A})$ is true in any interpretation I, so that $((\forall x_i)\, \mathcal{A} \rightarrow \mathcal{A})$ is logically valid.

(b) and (c) : See p 75, Hamilton.

Proposition 4.5.4. (Soundness theorem for K) :

For any wf $\mathcal{A}$ of $\mathcal{L}$, if $\vdash_{K} \mathcal{A}$ then $\mathcal{A}$ is logically valid.

Proof :

Note that $\vdash_{K} \mathcal{A}$ means that $\mathcal{A}$ is a theorem of K. In the proof of $\mathcal{A}$ there is a sequence $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n$ of $\mathcal{L}$ such that for each $i (1 \le i \le n)$ either $\mathcal{A}_i$ is an axiom of K or $\mathcal{A}_i$ follows from previous members of the sequence by MP and generalisation.

We use induction on n, i.e. on the number of members in the sequence of wfs of a proof of $\mathcal{A}$.

If $n = 1$, i.e., if the sequence of wfs in the proof contains only one member, namely $\mathcal{A}$, then $\mathcal{A}$ must be an axiom. But we know that any axiom is logically valid (by Propositions 4.1.3., 4.1.4.). So the theorem $\mathcal{A}$ is logically valid.

Suppose there are $n (> 1)$ members in the sequence of wfs in a proof $\mathcal{A}$. Suppose all theorems of K having proofs with shorter $(< n)$ sequences of wfs are logically valid.

If $\mathcal{A}$ follows from $\mathcal{B}$ and $\mathcal{B} \rightarrow \mathcal{A}$ by M.P., then both $\mathcal{B}$ and $\mathcal{B} \rightarrow \mathcal{A}$ have shorter sequences of wfs in their proof. So $\mathcal{B}$ and $(\mathcal{B} \rightarrow \mathcal{A})$ are logically valid. Then (by Proposition 3.4.5.) $\mathcal{A}$ is logically valid.

If $\mathcal{A}$ is deduced by Generalisation from $(\forall x_i)\, \mathcal{A}$, then $(\forall x_i)\, \mathcal{A}$ may be regarded as a theorem having a shorter sequence of wfs in its proof. So $(\forall x_i)\mathcal{A}$ is logically valid. Then (by Proposition 3.4.6.) $\mathcal{A}$ is logically valid.

Proposition 4.5.5. (Deduction Theorem for K) :

Let $\mathcal{A}$ and $\mathcal{B}$ be wfs of $\mathcal{L}$ and let Γ be a set (possibly empty) of wfs of $\mathcal{L}$. If $\Gamma \cup \{\mathcal{A}\} \vdash_K \mathcal{B}$, and **the deduction contains no application of generalisation involving a variable which occurs free in $\mathcal{A}$, then**

$$\Gamma \vdash_K (\mathcal{A} \rightarrow \mathcal{B})$$

Proof :

Let n be the number of wfs in the sequence constituting the deductiion of $\mathcal{B}$ from $\Gamma \cup \{\mathcal{A}\}$. We use induction on n.

When n = 1, then the sequence contains only $\mathcal{B}$. Therefore $\mathcal{B}$ is an axiom, or $\mathcal{A}$, or a member of Γ. We deduce that $\Gamma \vdash_K (\mathcal{A} \rightarrow \mathcal{B})$ in exactly the same way as we did in the corresponding proof for the Deduction Theorem for L.

Now let n > 1. Suppose that if $\mathcal{F}$ is a wf of $\mathcal{L}$ which can be deduced from $\Gamma \cup \{\mathcal{A}\}$ without using generalisation applied to a free variable of $\mathcal{A}$, in a deduction containing fewer than n wfs then $\Gamma \vdash_K (\mathcal{A} \rightarrow \mathcal{F})$

Case I :

$\mathcal{B}$ follows from previous wfs in the deduction by MP. The proof here is again the same as for L.

Case II :

$\mathcal{B}$ is an axiom, or $\mathcal{A}$, or a member of Γ. Again the proof is same as for L.

Case III :

$\mathcal{B}$ follows from previous wfs in the deduction by generalisation. So, suppose that $\mathcal{B}$ is $\mathcal{C}$ where $\mathcal{C}$ appears previously in the deduction. Thus $\Gamma \cup (\mathcal{A}) \vdash_K \mathcal{C}$,

and the deduction contains fewer than n wfs. So $\Gamma \vdash_K (\mathcal{A} \rightarrow \mathcal{C})$, since there is no application of Generalisation involving a free variable of $\mathcal{C}$. Also x_i cannot occur free in $\mathcal{A}$, as it is involved in an application of Generalisation in the deduction of $\mathcal{B}$ from $\Gamma \cup \{\mathcal{A}\}$. So we have a deduction of $(\mathcal{A} \rightarrow \mathcal{B})$ from Γ as follows :

(1)

.

. deduction of $(\mathcal{A} \to \mathcal{C})$ from Γ

.

.

(k) $(\mathcal{A} \to \mathcal{C})$

(k + 1) $(\forall x_i)(\mathcal{A} \to \mathcal{C})$ from (k) and Generalisation

(k + 2) $(\forall x_i)(\mathcal{A} \to \mathcal{C}) \to (\mathcal{A} \to (\forall x_i)\mathcal{C})$ (K6)

(k + 3) $(\mathcal{A} \to (\forall x_i)\mathcal{C})$ from (k + 1), (k + 2) and M.P.

So, $\Gamma \vdash_K (\mathcal{A} \to \mathcal{B})$ as required, and this concludes the inductive proof.

Corollaries :

(1) If $\Gamma \cup \{\mathcal{A}\} \vdash_K \mathcal{B}$, and $\mathcal{A}$ is a closed wf, then $\Gamma \vdash_K (\mathcal{A} \to \mathcal{B})$.

Proof :

Immediate. "$\mathcal{A}$ is closed" means "there is no variable occurring free in $\mathcal{A}$", and therefore the deduction does not contain application of Generalisation involving a variable which occurs free in $\mathcal{A}$.

(2) For any wfs $\mathcal{A}, \mathcal{B}, \mathcal{C}$ of $\mathcal{L}$,

$$\{(\mathcal{A} \to \mathcal{B}), (\mathcal{B} \to \mathcal{C})\} \vdash_K (\mathcal{A} \to \mathcal{C})$$

Proof :

The proof is identical with the proof of a similar result in propositional calculus. (The **Hypothetical Syllogism, H,S.**)

Proposition 4.5.6. :

Suppose $\mathcal{A}$ and $\mathcal{B}$ are wfs of $\mathcal{L}$, Γ a set of wfs of $\mathcal{L}$ and $\Gamma \vdash_K (\mathcal{A} \to \mathcal{B})$. Then

$$\Gamma \cup (\mathcal{A}) \vdash_K \mathcal{B}$$

Proof :

Identical with the corresponding proposition in propositional calculus.

Example 4.5.1. :

If x_i does not occur free in $\mathcal{A}$, then

$$\vdash_K ((\mathcal{A} \to (\forall x_i)\mathcal{B}) \to (\forall x_i)(\mathcal{A} \to \mathcal{B}))$$

Proof :

(1) $(\mathcal{A} \rightarrow (\forall x_i)\,\mathcal{B})$ assumption

(2) $((\forall x_i)\,\mathcal{B} \rightarrow \mathcal{B})$ (K4)

(3) $(\mathcal{A} \rightarrow \mathcal{B})$ (1), (2) and H.S.

(4) $(\forall x_i)(\mathcal{A} \rightarrow \mathcal{B})$ (3) and Generalisation

So, $(\mathcal{A} \rightarrow (\forall x_i)\,\mathcal{B}) \vdash_{K} (\forall x_i)(\mathcal{A} \rightarrow \mathcal{B})$.

Now, Generalisation is used in the above deduction only by using the variable x_i which does not occur free in $(\mathcal{A} \rightarrow (\forall x_i)\,\mathcal{B})$. Hence we can apply the deduction theorem to obtain :

$$\vdash_{K} ((\mathcal{A} \rightarrow (\forall x_i)\,\mathcal{B}) \rightarrow (\forall x_i)(\mathcal{A} \rightarrow \mathcal{B}))$$

Proposition 4.5.7. (The Adequacy Theorem for $K_{\mathcal{L}}$) :

If $\mathcal{A}$ is a logically valid wf of $\mathcal{L}$, then $\mathcal{A}$ is a theorem of $K_{\mathcal{L}}$.

Proof :

See Hamilton, pages 99-100.

Summary :

- In a given interpretation I, a wf $\mathcal{A} \rightarrow \mathcal{B}$ is false if and only if $\mathcal{A}$ is true and $\mathcal{B}$ is false.
- If in a particular interpretation I, the wfs. $\mathcal{A}$ and $\mathcal{A} \rightarrow \mathcal{B}$ are true then $\mathcal{B}$ is also true.
- Let $\mathcal{A}$ be a wf of $\mathcal{L}$, and let I be an interpretation of $\mathcal{L}$. Then $I \models \mathcal{A}$ if and only if $I \models (\forall x_i)\,\mathcal{A}$ where x_i is any variable.
- A wf $\mathcal{A}$ of L is logically valid if and only if $(\forall x_i)\,\mathcal{A}$ is logically valid where x_i is any variable.
- If $\mathcal{A}$ is a wf of $\mathcal{L}$ and $\mathcal{A}$ is a tautology, then $\mathcal{A}$ is a theorem of K.
- All instances of (K1), (K2) and (K3) are logically valid.
- **Soundness theorem for K :**

  For any wf $\mathcal{A}$ of $\mathcal{L}$, if $\vdash_{K.} \mathcal{A}$ then $\mathcal{A}$ is logically valid.
- **(Deduction Theorem for K) :**

  Let $\mathcal{A}$ and $\mathcal{B}$ be wfs of $\mathcal{L}$ and let Γ be a set (possibly empty) of wfs of $\mathcal{L}$. If $\Gamma \cup \{\mathcal{A}\} \vdash_{K} \mathcal{B}$, and the deduction contains no application of generalisation involving a variable which occurs free in $\mathcal{A}$, then

  $$\Gamma \vdash_{K.} (\mathcal{A} \rightarrow \mathcal{B})$$

- **(The Adequacy Theorem for $K_{\mathcal{L}}$) :**

  If $\mathcal{A}$ is a logically valid wf of $\mathcal{L}$, then $\mathcal{A}$ is a theorem of $K_{\mathcal{L}}$.

G.U. Questions

1997

1. Give an interpretation so that the wf

$$(\forall x_1)(\forall x_2)(\sim(\forall x_3)(\sim A_1^2(f_1^2(x_1, x_3), x_2)))$$

can be translated to :

' For any two positive rational numbers x and y there is a positive rational number z such that xz = y'. 3

2. Justify the truth of the following statement.

'The truth or falsity of a statement representating a wf in first order language depends on the interpretation.' 6

3. Define the terms **substitution instance** and **tautology** as applied in predicate calculus.

Examine if the wf

$$(\forall x_1)A_1^1(x_1) \wedge (\forall x_2)A_1^2(x_1, x_2) \rightarrow (\forall x_1)A_1^1(x_1)$$

is a tautology of predicate calculus. $2+5=7$

1998

1. Express the following statements in first order language :

(i) Some real numbers are rational.

(ii) No woman is both a politician an a housewife. $2+2=4$

2. With usual symbolism of predicate calculus examine if the following expressions are well-formed formulas :

(i) $A_1^2(f_1^1(x_1), x_1)$

(ii) $\sim(\forall x_2)A_1^2(x_1, x_2)$

(iii) $\sim A_1^1(x_1) \rightarrow A_1^1(x_2)$ $1+1+1=3$

3. Consider the following well formed formula of predicate calculus :

$$A_1^2(f_1^2(x_1, x_2), f_2^2(x_1, x_2))$$

Give an interpretation to the above formula to explain the notion of **satisfaction** of a formula. 5

4. Prove that a well-formed wf $\mathcal{A}$ of predicate calculus is true in an interpretation I if and only if $(\forall x)\mathcal{A}$ where x is any variable is true in I. 6

1999

1. Describe a **First Order Language of Predicate Calculus** and define a well-formed in this language. $2+2=4$

2. Define **interpretation** of a wf of **Predicate Calculus**. Give an interpretation to the followings wf and examine its truth-value :

$$(\forall x_1)(\forall x_2)(\sim \forall x_3)(-A_1^2)(f_1^2(x_1, x_3), x_2) \qquad 2 + 5 = 7$$

3. In a given interpretation I, prove that $A \rightarrow B$ is false if and only if A is true and B is false. 5

4. Let A be a wf of Predicate calculus and I an interpretation. Prove that $I \models A$ if and only if $I \models (\forall x_1) A$ where x_1 is any variable. 6

1997

1. State a set of **axiom schemas** and **rules of deduction** for a formal theory K of predicate calculus. Prove that any tautology of propositional calculus is a theorem of predicate calculus.
Is the converse true? 3 + 4 + 1 = 8

2. Prove **Deduction Theorm of K** :

Let $\mathcal{A}$ and $\mathcal{B}$ be wfs of propositional calculus L and let Γ be a set(possibly empty) of wfs of L. If

$$\Gamma \cup \{\mathcal{A}\} \vdash_K \mathcal{B}$$

and the deduction contains no application of Generalization involving a variable which occurs free in $\mathcal{A}$, then

$$\Gamma \vdash_K (\mathcal{A} \rightarrow \mathcal{B})$$

1998

1. If $\mathcal{A}$ is a theorem in a formal theory of predicate calculus, then prove that $\mathcal{A}$ is logically valid 6

2. If $\mathcal{S}$ is a consistent first order system with equality, then prove that the following is a theorem of $\mathcal{S}$:

$$(\forall x_1)(\forall x_2)(A_1^2(x_1, x_2) \rightarrow A_1^2(x_2, x_1)) \qquad 5$$

3. Describe a first order system using a formal language which contains no individual constants.?

1999

1. Present predicate calculus as an Axiomatic Theory K stating its axiom schemas and the rules of inference.
Prove that a tautology of Propositional Calculus is a theorem in K.

4 + 6 = 10

●●●

UNIT : 5
Mathematical systems

Introduction

Units 1 to 4 are not mathematics. The systems L and $K_{\mathcal{L}}$ are systems of logical deduction. We have had to use some mathematical techniques in order to obtain proofs of propositions, but those techniques have been of an elementary nature, principally properties of natural numbers. The mathematician interested in the foundations of his subject seeks to clarify the assumptions he makes and the procedures he uses. We can use the system $K_{\mathcal{L}}$ in such a clarification. $K_{\mathcal{L}}$ embodies procedures of logical deducation as used by mathematicians. We have seen that the absence of restrictions on the language $\mathcal{L}$ make our results about K very general, and that the symbols of a given $\mathcal{L}$ can be interpreted in many different ways. For any $\mathcal{L}$, however, there is a class of *wfs*. whose truth does not depend on the interpretation of the symbols, namely the class of logically valid *wfs*., i.e. the class of theorems of $K_{\mathcal{L}}$. If $\mathcal{L}$ is interpreted in a mathematical way, as it is in our examples, the theorems of $K_{\mathcal{L}}$ are interpreted as mathematical truths. They are mathematical statements which are true because of their logical structure rather than because of their mathematical content. For example, in the arithmetic interpretation N, the *wf.*

$$(\forall x_1)(\forall x_2)(A_1^2(x_1, x_2) \rightarrow A_1^2(x_1, x_2))$$

which is logically valid, is interpreted as a mathematical statement, namely: 'for all natural numbers x and y, if $x = y$ then $x = y$', which is true by virtue of its logical structure. On the other hand, the *wf.*

$$(\forall x_1)(\forall x_2)(A_1^2(x_1, x_2) \rightarrow A_1^2(x_2, x_1))$$

is interpreted as the mathematical statement: 'for all natural numbers x and y, if $x = y$ then $y = x$', which is true. That it is true is, however, a consequence of the meaning of '=' rather than merely its logical structure. Indeed this *wf.* is not logically valid. It is not difficult to find an interpretation in which A_1^2 is not interpreted as =, in which it is false. It follows that this *wf.* is not a theorem of $K_{\mathcal{L}}$. Thus the theorems of $K_{\mathcal{L}}$ have in themselves no mathematical value. Each of our mathematical formal systems will be an extension of some $K_{\mathcal{L}}$ obtained by including appropriate additional axioms so that the theorems of the system will represent mathematical truths as well as logical truths. If our formal system is to be a mathematical system then it is clearly desirable to have as theorems all *wfs*. whose interpretations are mathematical truths (or, if that is not possible, as many such *wfs*. as we can).

What constitutes a mathmatical truth depends to a very larg extent on the mathematical context. For example, the statement

$$(\forall x)(\forall y)(xy = yx)$$

is true when regarded as a statement about natural numbers, but is not necessarily true when regarded as a statement about elements f an arbitrary group. We shall show, by means of examples, how different mathematical contexts can be represented by different formal systems, so that, in particular, the above statement would be the interpretation of a theorem of formal arithmetic but the interpretation of a non-theorem of formal group theory. The context will determine the language $\mathcal{L}$ (as in the case of arithmetic) and it will also determine a set of *proper axioms*. The word 'proper' is used in order to distinguish these from (*K* 1)–(*K*6), which are *logical axioms*, and are common to all our systems. Having specified $\mathcal{L}$, the proper axioms are *wfs.* of $\mathcal{L}$ which, when added as new axioms, give an extension of $K_{\mathcal{L}}$ in which mathematical truths of the particular context (as well as logical truths) appear as interpretations of theorems.

5.1. First order systems with equality

Mathematics can very rarely do without the relation of equality. The symbol '=' does not appear in our formal languages, but we have used it in examples as the interpretation of the predicate symbol A_1^2. In all our examples of mathematical systems we shall include A_1^2 in the language, and = will be its intended interpretation.

As we observed above, the *wf.* $(\forall x_1)(\forall x_2)(A_1^2(x_1, x_2) \to A_1^2(x_2, x_1))$ is not a theorem of $K_{\mathcal{L}}$ but we would like it to be a theorem of our mathematical extensions of $K_{\mathcal{L}}$. One way of ensuring this would be to include it among the proper axioms of each mathematical system. But there are clearly other *wfs.* which would require similar treatment, for example $(\forall x_1) A_1^2(x_1, x_1)$. We do not need to include as axioms all such *wfs.*, but we take as *axioms for equality* a set of them from which others may be deduced.

$(E7)\, A_1^2(x_1, x_1)$.

$(E8)\, A_1^2(t_k, u) \to A_1^2\,(f_1^n(t_1, \ldots t_k \ldots, t_n),\ \ f_1^n(t_1, \ldots u \ldots, t_n)$, where $t_1, \ldots, t_n$, u are any terms, and f_1^n is any predicate symbol of $\mathcal{L}$.

$(E9)\, A_1^2(t_k, u) \to A_1^n\,(t_1, \ldots t_k \ldots, t_n) \to A_1^n\,(t_1, \ldots u \ldots, t_n)$ where $t_1, \ldots t_n$, u are any terms, and A_1^n is any predicate symbol of $\mathcal{L}$.

***Notes* 5.1.1**

(*a*) (*E*8) and (*E*9) are axiom schemes each representing a number of axioms, possibly infinitely many, depending on the number of function letters and predicate symbols in $\mathcal{L}$.

(*b*) All of these axioms have free variables occurring in them. They have been written in this way for the sake of clarity and for ease of application later. We know, however, that for any *wf.* $\tilde{A}$, whose universal closure is $\tilde{A}'$, $\tilde{A} \vdash \tilde{A}'$ and $\tilde{A}' \vdash \tilde{A}$ so an equivalent set of axioms would be the universal closures of these.

(*c*) As a consequence of (*b*), and Proposition 4.18, regarding change of bound variables, the fact that the variable x_1 in particular appears in (*E7*) has no significance. For example, $A_1^2(x_5, x_5)$ is a consequence of (*E7*), by means of the deduction :

(1)	$A_1^2(x_1, x_1)$	(*E7*)
(2)	$(\forall x_1) A_1^2(x_1, x_1)$	(1) Generalisation
(3)	$(\forall x_5) A_1^2(x_5, x_5)$	(2), Proposition 4.18
(4)	$(\forall x_5) A_1^2(x_5, x_5) \to A_1^2(x_5, x_5)$	(*K5*)
(5)	$A_1^2(x_5, x_5)$	(3), (4) *MP*.

All mathematical systems which we describe will be extensions of $K_{\mathcal{L}}$ (for some $\mathcal{L}$) which include amongst their axioms ($E7$) and all appropriate (depending on $\mathcal{L}$) instances of ($E8$) and ($E9$).

***Remark* 5.1.2**

The need to include ($E7$) should be clear. It ensure that in any model the interpretation of A_1^2 behave in one respect like =. ($E8$) and ($E9$) are more complex, but their inclusion ensures that in any model, the interpretation of A_1^2 behaves like = in another respect, namely that equals may be substituted for one another.

***Definition* 5.1.3**

The axioms ($E7$), ($E8$) and ($E9$) are called *axioms for equality*. Any extension of $K_{=}$ which includes amongst its axioms ($E7$) and all appropriate instances of ($E8$) and ($E9$) is called a *first order system with equality.*

***Proposition* 5.1.4**

Let S be a first order system with equality. Then the following are theorems of S.

(i) $(\forall x_1)A_1^2(x_1,x_1)$

(ii) $(\forall x_1)(\forall x_2)(A_1^2(x_1,x_2) \rightarrow A_1^2(x_2,x_1))$,

(iii) $(\forall x_1)(\forall x_2)(\forall x_3)(A_1^2(x_1,x_2) \rightarrow (A_1^2(x_2,x_3) \rightarrow A_1^2(x_1,x_3)))$

Proof. (i) Immediate, by Generalisation, from ($E7$).

(ii) We give a proof in S.

(1) $A_1^2(x_1,x_2) \rightarrow (A_1^2(x_1,x_1) \rightarrow A_1^2(x_2,x_1))$ ($E9$)

(2) $A_1^2(x_1,x_2) \rightarrow (A_1^2(x_1,x_1) \rightarrow A_1^2(x_2,x_1))) \rightarrow$
$((A_1^2(x_1,x_2) \rightarrow A_1^2(x_1,x_1)) \rightarrow (A_1^2(x_1,x_2) \rightarrow A_1^2(x_2,x_1)))$ ($K2$)

(3) $(A_1^2(x_1,x_2) \rightarrow A_1^2(x_1,x_1)) \rightarrow (A_1^2(x_1,x_2) \rightarrow A_1^2(x_2,x_1))$ (1), (2) MP

(4) $(A_1^2(x_1,x_1) \rightarrow A_1^2(x_1,x_2) \rightarrow (A_1^2(x_1,x_1)))$ ($K1$)

(5) $A_1^2(x_1,x_1)$ ($E7$)

(6) $(A_1^2(x_1,x_2) \rightarrow A_1^2(x_1,x_1))$ (4), (5), MP

(7) $(A_1^2(x_1,x_2) \rightarrow A_1^2(x_1,x_1))$ (3), (6), MP

(8) $(\forall x_1)(\forall x_2)(A_1^2(x_1,x_2) \rightarrow A_1^2(x_2,x_1))$ (7), Generalisation

(iii) Again we give a proof in S.

(1) $(A_1^2(x_2,x_1) \rightarrow (A_1^2(x_2,x_3) \rightarrow (A_1^2(x_1,x_3)))$ ($E9$)

(2) $(A_1^2(x_1,x_2) \rightarrow (A_1^2(x_2,x_1))$ (ii) above

(3) $(A_1^2(x_1,x_2) \rightarrow (A_1^2(x_2,x_3) \rightarrow A_1^2(x_1,x_3)))$ (1), (2), HS

(4) $(\forall x_1)(\forall x_2)(\forall x_3)(A_1^2(x_1,x_2) \rightarrow (A_1^2(x_2,x_3) \rightarrow A_1^2(x_1,x_3)))$ (3), Generalisation

Δ Thus, since each of (i), (ii), (iii) in the above proposition must be true in any model of S, the symbol A_1^2 will be interpreted in any model by a relation which is reflexive, symmetric and transitive, i.e. an equivalence relation. Now = is the intended interpretation for A_1^2. In an arbitrary interpretation, the axioms could well be false, so A_1^2 could be interpreted by any binary relation, but in a model of S we have seen that the axioms must be true and, as above, A_1^2 must be interpreted as an equivalence relation. However, the axioms ($E7$), ($E8$) and ($E9$) do not ensure that in any model of S the interpretation of A_1^2 is actually =.

***Example* 5.1.5**

Consider the first order language L with variables $x_1, x_2, \ldots$, function letter f_1^2, and predicate letter A_1^2. Define an interpretation I as follows. D_I is the set Z of all integers, $f_1^{-2}(x, y)$ is $x+y$, and $A_1^{-2}(x, y)$ holds if and only if $x \equiv y \pmod 2$, for $x, y \in$ Z. The axioms for equality are true in this interpretation.
For ($E7$), its interpretation is $x \equiv x \pmod 2$, which is true,
For ($E8$), as a particular case, consider

$$A_1^2(x_1, y_2) \rightarrow A_1^2(f_1^2(x_1, x_3), f_1^2(x_2, x_3)).$$

This is interpreted as :

$$\text{if } x \equiv y \pmod 2 \text{ then } x + z \equiv y + z \pmod 2$$

which is true. Verification of ($E8$) in its full generality is left as an exercise.
For ($E9$), there are only two instances to be verified, since L contains only one predicate letter. These are

$$(A_1^2(t, u) \rightarrow (A_1^2(t, v) \rightarrow A_1^2(u, v))).$$

and

$$(A_1^2(t, u) \rightarrow (A_1^2(v, t) \rightarrow A_1^2(v, u))).$$

The interpretations of these are, respectively.
if $x \equiv y \pmod 2$ then $x \equiv z \pmod 2$ implies $y \equiv z \pmod 2$
and
if $x \equiv y \pmod 2$ then $z \equiv x \pmod 2$ implies $z \equiv y \pmod 2$, which are true.

$\triangleright$ This example shows that in a model of ($E7$), ($E8$) and ($E9$) the symbol A_1^2 need not necessarily be interpreted by =. However, the following proposition restores the situation.

***Proposition* 5.1.6**

If S is a consistent first order system with equality, then S *has a model in which the interpretation of* A_1^2 is =.

Proof. By Proposition 4.44, if S is consistent, then S has a model, M, say. A_1^{-2} is an equivalence relation on D_M because of Proposition 5.4.

Denote the equivalence class containing x by $[x]$. Now define a new interpretation M^* as follows. The domain of M^* is $\{[x]: x \in D_M\}$, a_i is interpreted by $[\bar{a}_i]$, for each i, f_i^n is interpreted by $\hat{f}_i^n$ where, for $y_1, \ldots, y_n \in D_M$

$$f_i^n([y_1], \ldots, [y_n]) = [\hat{f}_i^n(y_1, \ldots, y_n)],$$

and A_i^n is interpreted by $\hat{A}_i^n$, where, for $y_1, \ldots y_n \in D_M$

$\hat{A}_i^n([y_1], \ldots, [y_n])$

holds if and only if $\bar{A}_i^n(y_1, \ldots, y_n)$ holds, where $\bar{a}_i$, $\bar{f}_i^n$, $\bar{A}_i^n$ are the interpretations of the symbols of $\propto$ in M. It is a lengthy but not difficult task to verify that these are well defined and that M^* is a model of S. For example, let f be a one-place function letter of $\propto$ and $\bar{f}$ its interpretation in M. Suppose that a and b are members of D_M and that $[a] = [b]$. We have to show that $[\bar{f}(a)] = [\bar{f}(b)]$.

Now

$$\vdash (A_1^2(x_1, x_2) \to A_1^2(f(x_1), f(x_2))). \qquad (E8)$$

Hence $(A_1^2(x_1, x_2) \to A_1^2(f(x_1), f(x_2)))$. is true in M, since M is a model, and so $\bar{A}_1^2(a,b)$ implies $\bar{A}_1^2(f(a), f(b))$, i.e. $[a] = [b]$ implies $[\bar{f}(a)] = [\bar{f}(b)]$.

Also the interpretation of A_1^2 in M^* is $=$, since $\hat{A}_1^2([x], [y])$ holds if and only if $\bar{A}_1^2(x, y)$ holds. i.e. and only if $[x] = [y]$.

▷ This proof can be well illustrated by our last example in which we gave a model where A_1^2 was not interpreted as $=$. In that example we had $\bar{A}_1^2(x, y)$ if and only if $x \equiv y \pmod 2$ (x and y integers). Define a new model, with domain $\{[0], [1]\}$, in which f_1^2 and A_1^2 are interpreted by $\hat{f}_1^2$ and $\hat{A}_1^2$, given by

$\hat{f}_1^2([x], [y]) = [\bar{f}_1^2(x, y)] = [x + y]$,

$\hat{A}_1^2([x], [y])$ holds if and only if $\bar{A}_1^2(x, y)$ holds,

i.e. if and only if $x \equiv y \pmod 2$

i.e. if and only if $[x] = [y]$.

***Definition* 5.1.7**

Let S be a first order system with equality. A *normal* model of S is a model in which A_1^2 is interpreted as $=$. We shall be concerned in what follows mostly with normal models, since they represent the intended mathematical situation regarding the interpretation of A_1^2.

Note. Of course it does not matter that we have chosen A_1^2 to stand for equality. We could have chosen, say A_{17}^2 in which case axioms (*E7*), (*E8*) and (*E9*) would have involved this predicate symbol instead of A_1^2.

Δ For the rest of this chapter we shall be dealing with first order systems with equality in which A_1^2 stands for equality. The proof of Proposition 5.4 indicates how repetitive writing out proofs can become, and we can alleviate this some what by introducing the symbol = into our language in place of A_1^2.

Notation. Write $t_1 = t_2$ in place of $A_1^2(t_1, t_2)$ where t_1 and t_2 are terms of $\mathcal{L}$.

Axioms (*E7*), (*E8*) and (*E9*) can now be written in a simplified form, and in way which makes their significance much clearer.

(*E7'*) $x_1 = x_1$

(*E8'*) $(t_k = u \rightarrow (f_i^n(t_1, ..., t_k, ..., t_n) = f_i^n(t_1, ..., u, ..., t_n)))$.

$t_1, ..., t_n, u, f_i^n$ as in (*E8*)

(*E9'*) $(t_k = u \rightarrow (A_i^n(t_1, ..., t_k, ..., t_n) \rightarrow A_i^n(t_1, ..., u, ... t_n)))$,

$t_1, ..., t_n, u, A_i^n$ as in (*E9*)

∀

The symbol = is not the only one which we have introduced into the formal language in addition to the original alphabet of symbols. For example we use $(\exists x_i)$ as an abbreviation for $\sim(\forall x_i)\sim$, and we use $(\mathcal{A} \leftrightarrow \mathcal{B})$ as an abbreviation for $\sim((\mathcal{A} \rightarrow \mathcal{B}) \rightarrow \sim(\mathcal{B} \rightarrow \mathcal{A}))$. It is sometimes convenient to write $(\mathcal{A} \vee \mathcal{B})$ as an abbreviation for $(\sim\mathcal{A} \rightarrow \mathcal{B})$, and $(\mathcal{A} \wedge \mathcal{B})$ for $\sim(\mathcal{A} \rightarrow \sim\mathcal{B})$. These correspond with our intuitive ideas from Chapter 1, and the use of these new symbols clearly does not extend our formal systems in any way. It is a convenience to avoid lengthy repetition of symbols. In the different contexts which we are about to describe it is possible and sometimes desirable to extend the practice, and introduce further *defined symbols*. There is one useful one in particular which applies in any first order system with equality. This is the symbol for 'there exists a unique.... such that'.

Notation. $(\exists_1 x_i)\mathcal{A}(x_i)$ is an abbreviation for the formula

$$(\exists x_i)(\mathcal{A}(x_i) \wedge (\forall x_j)(\mathcal{A}(x_j) \rightarrow x_i = x_j)).$$

5.2 The theory of groups

Group theory is perhaps the most familiar branch of mathematics which is based explicitly on a simple set of axioms, so let us use this 'mathematical context' to illustrate how mathematical systems arise as extensions of $K_{\mathcal{L}}$.

First we must describe an appropriate first order language $\mathcal{L}$, so let $\mathcal{L}_G$ be the first order language with the following alphabet of symbols:

variables $x_1, x_2, \ldots$

individual constant a_1 (identity)

function symbols f_1^1, f_1^2 (inverse, product)

predicate symbol =

punctuation (,) , ,

logical symbols $\forall, \sim, \rightarrow$,

Now define $\mathcal{G}$ to be the extension of $K\mathcal{L}_G$ whose proper axioms are ($E7$), all appropriate instances of ($E8$) and ($E9$), and the following:

($G1$) $f_1^2(f_1^2(x_1, x_2), x_3) = f_1^2(x_1, f_1^2(x_2, x_3))$ (associative law)

($G2$) $f_1^2(a_1, x_1) = x_1$ (left identity)

($G3$) $f_1^2(f_1^1(x_1), x_1) = a_1$ (left inverse).

As previously, it does not matter whether universal quantifiers are included in these axioms for each free variable or not. An equivalent set of axioms would be the universal closures of these.

($G1$), ($G2$) and ($G3$) are merely translations of the usual group axioms. Normally ($G2$) and ($G3$) are stated in the form 'There exists a left identity' and 'For each element there exists a left inverst'. Our axioms here do not explicitly assert existence. They merely state that a_1 and $f_1^1(x_1)$. when interpreted in a model, must have the appropriate properties. To assert existence is unnecessary since in any model of this system there will be interpretations of a_1 and of f_1^1, and so the identity and inverse will automatically exist. Similarly, the group axiom concerning closure under the group operation is unnecessary here because the interpretation of f_1^2 in a model is necessarily a two place function with values in the domain of the model.

Given such a system of group theory, we can convert any standard proof from an algebra textbook of a result about elements of groups into a formal proof in the system. Such a procedure would have little practical use, for a formal proof in $\mathcal{G}$ is necessarily rather complicated, and the large numbers of purely manipulative steps would obscure the intuitive ideas involved, as the following example shows.

***Example* 5.2.1**

In any group G with identity element e, $e(ee) = e$. Corresponding to this, let us give a formal proof in the system $\mathcal{G}$ of the *wf.*

$$f_1^2(a_1, f_1^2(a_1, a_1)) = a_1.$$

(1) $f_1^2(a_1, x_1) = x_1$

(2) $(\forall x_1)(f_1^2(a_1,x_1)=x_1)$ (1), Generalisation

(3) $(\forall x_1)(f_1^2(a_1,x_1)=x_1)\to(f_1^2(a_1,a_1)=a_1)$ (K5)

(4) $f_1^2(a_1,a_1)=a_1$ (2), (3), *MP*

(5) $(\forall x_1)(f_1^2(a_1,x_1)=x_1)\to(f_1^2(a_1,f_1^2(a_1,a_1))$
$=f_1^2(a_1,a_1))$ (K5)

(6) $f_1^2(a_1,f_1^2(a_1,a_1))=f_1^2(a_1,a_1)$ (2), (5), *MP*

(7) $(f_1^2(a_1,a_1)=a_1)\to(f_1^2(a_1,f_1^2(a_1,a_1))$
$=f_1^2(a_1,a_1)\to f_1^2(a_1,f_1^2(a_1,a_1))=a_1)$ (E9)

(8) $(f_1^2(a_1,f_1^2(a_1,a_1))=f_1^2(a_1,a_1)$
$\to f_1^2(a_1,f_1^2(a_1,a_1))=a_1)$ (4), (7), *MP*

(9) $f_1^2(a_1,f_1^2(a_1,a_1))=a_1$ (6), (8), *MP*

In comparison with this, a standard proof of $e(ee)=e$ for any group is a triviality. More complicated results about groups are reflected in still more complicated formal proofs in $\mathcal{G}$. Particular examples are not very rewarding, but a further idea of the complications involved will be obtained by attempting to prove in ∝ the *wf.*

$$f_1^2(x_1,a_1)=x_1$$

Corresponding to the property of groups that the left identity is also a right identity.

▷ It should be clear that any group G is a model of the system $\mathcal{G}$ provided that a_1 is interpreted as the identity element of G. f_1^1 as the inverse, f_1^2 as the group operation and = as equals. However, there are other models, as we shall see.

***Example* 5.2.2**

Construct an interpretation I of the system $\mathcal{G}$ as follows. Let D_I be the set Z of integers, let a_1 be interpreted as 0, let

$$\bar{f}_1^1(x)=-x \text{ for } x\in \mathbb{Z},$$

and

$$\bar{f}_1^2(x,y)=x+y \text{ for } x,y\in \mathbb{Z}.$$

and let = interpreted by congruence (mod m), where m is some fixed positive integer. (Although we are using = as a symbol of $L_{\mathcal{G}}$, as we have seen above, it need not be interpreted always as actual equality.) I is a model of $\mathcal{G}$. To verify this we must show that every axiom of $\propto$ is true in I. That $(K1)$–$(K6)$ are true requires no verification since they are logically valid. That $(E7)$, $(E8)$ and $(E9)$ are true is verified just as in Example 5.5. Let us look more closely at $(G1)$, $(G2)$ and $(G3)$.

$(G1)$ is interpreted as

$$(x + y) + z \equiv x + (y + z) \pmod{m}.$$

$(G2)$ is interpreted as

$$0 + x \equiv x \pmod{m}.$$

$(G3)$ is interpreted as

$$-x + x \equiv 0 \pmod{m}.$$

All of these are true statements, for any $x, y, z \in \mathbb{Z}$. Thus I is a model of $\mathcal{G}$. However, I is not a group. Indeed it involves the extraneous relation of congruence. However, the reader with some experience of group theory or number theory will realise that there is a group in the background waiting to be discovered. From the model I we can construct a normal model I^* by the procedure of Proposition 5.6. The domain of I^* is the set of congruence classes of integers (mod m), a_1 is interpreted by 0_m (the class containing 0), f_1^2 is interpreted by $+$ (which is well defined on congruence classes), f_1^1 is interpreted by 'additive inverse' (which is again well-defined), and $=$ is interpreted by equals. I^* is a normal model, and it is a group.

▷ In general, any group is a normal model for the formal system of group theory, and conversely any normal model of the system is a group. So to make mathematical sense of the system we must restrict our attention to normal models. It is unfortunate, perhaps, but it is impossible to give axioms for equality which force the interpretation to be actual equality. It will always be possible to construct a model in which $=$ is interpreted by some other equivalence relation.

The reason for constructing this formal system of group theory is not to provide any shortcuts or new methods for obtaining results about groups and their elements. As we have seen, the methods of proof within $\mathcal{G}$ are so unwieldy as to be useless for this purpose. What we have gained by describing the system $\mathcal{G}$ is that we have made precise and explicit all the assumptions and procedures which mathematicians use in the context of group theory, including the logical ones as well as the mathematical ones. In this way we have clarified this part of mathematics.

Groups have been treated in detail, and similar treatment can be given for other sorts of abstract algebraic systems, for example rings, fields, vector spaces, lattices, boolean algebras, etc. Each of these is known to be characterised by a finite set of axioms, and these can be easily translated into an appropriate formal language. Indeed every area of mathematics which is characterised by a set of axioms may be treated in a similar way. For example, Euclidean geometry can be based on a rather lengthy and complex set of axioms, and a formal system would have to include predicate letters intended to be interpreted by 'is a point', 'is a line', 'interesect' etc. Also an axiomatic system for the real numbers can be described, by means of the axioms for a complete ordered field.

There are two areas of mathematics which are particularly important when treated in this way. They are arithmetic and set theory. Each would require a whole book for a full treatment, but we shall merely try to explain why they have a special position. It is only within the framework of an explicit formal system that questions of consistency or of the relationships between different assumptions or of the position and use of fundamental assumptions can be clarified. Set theory serves as a foundation for all of mathematics, so its logical base is of over-riding importance. Arithmetic is an elementary fragment of mathematics, and its significance lies in the methods used to show that the search for a formal system which would enable any mathematical proposition to be tested must be fruitless. Any mathematical system in which ordinary arithmetic can be performed cannot be such a universal system, for the set of theorems of any consistent extension of arithmetic (in a sense which will be made precise) omits at least one true proposition. Some systems which are extensions of arithmetic (e.g. the theory of groups) do not have this property. However a system which includes mathematical analysis or is intended to embrace mathematics as a whole will certainly include arithmetic, and so will suffer from this shortcoming.

5.3 First order arithmetic

We develop the ideas involved in the arithmetic interpretation N first introduced in Chapter 3. The language $\mathcal{L}_N$ we take to include variables $x_1, x_2, \ldots$, the individual constant a_1 (for 0), the function letters f_1^1, f_1^2, f_2^2 (successor, sum and product), and the predicate symbol =, as well as punctuation, connectives and quantifier. Let us denote by $\mathcal{N}$ the first order system which is the extension of $K_{\mathcal{L}N}$, obtained by including as additional axioms *(E7)*, all appropriate instances of *(E8)* and *(E9)*, and the following six axioms and one axiom scheme.

(*N*1) $(\forall x_1) \sim (f_1^1(x_1) = a_1)$.

(*N*2) $(\forall x_1),(\forall x_2)(f_1^1(x_1) = f_1^1(x_2) \rightarrow x_1 = x_2)$

(*N*3) $(\forall x_1)\ (f_1^2(x_1, a_1) = x_1)$

(*N*4) $(\forall x_1)\ (\forall x_2)(f_1^2(x_1, f_1^1(x_2)) = f_1^1(f_1^2(x_1, x_2)))$.

(*N*5) $(\forall x_1)\ (f_2^2(x_1, a_1) = a_1)$

(*N*6) $(\forall x_1)\ (\forall x_2)(f_2^2(x_1, f_1^1(x_2)) = f_1^2(f_2^2(x_1, x_2), x_1))$.

(*N*7) $\mathcal{A}(a_1) \rightarrow ((\forall x_1)(\mathcal{A}(x_1) \rightarrow \mathcal{A}(f_1^1(x_1))) \rightarrow (\forall x_1)\mathcal{A}(x_1))$

for each *wf.* $\mathcal{A}(x_1)$ of $\mathcal{L}_N$ in which x_1 occurs free.

Notation. As yet we cannot know whether, for example, f_1^2 must, in any normal model, be imerpreted as addition (or a function with the same properties as the sum function), but it will make the system $\mathcal{N}$ much clearer and the axioms above much easier to understand if we modify $\mathcal{L}_N$ immediately by using the symbols +, × and instead of f_1^2, f_2^2 and f_1^1 respectively. To be explicit, we shall write

$$t_1 + t_2 \text{ for } f_1^2(t_1, t_2)$$

$$t_1 + t_2 \text{ for } f_2^2(t_1, t_2)$$

and

$$t' \text{ for } f_1^1(t)$$

where t, t_1, t_2 are any terms. Also we shall use the symbol 0 rather than a_1. The dangers of doing this must be emphasised once again. Having done it, we must not assume that these new symbols are necessarily always interpreted by the functions or objects which they normally represent.

Using these symbols, the axioms (N1)-(N7) may be rewritten as follows

(N1*) $\quad (\forall x_1) \sim (x_1' = 0)$,

(N2*) $\quad (\forall x_1)(\forall x_2)(x_1' = x_2' \to x_1 = x_2)$

(N3*) $\quad (\forall x_1)(x_1 + 0 = x_1)$,

(N4*) $\quad (\forall x_1)(\forall x_2)(x_1 + x_2' = (x_1 + x_2)')$.

(N5*) $\quad (\forall x_1)(x_1 \times 0 = 0)$,

(N6*) $\quad (\forall x_1)(\forall x_2)(x_1 \times x_2' = (x_1 \times x_2) + x_1)$.

(N7*) $\quad \mathcal{A}(0) \to (\forall x_1)(\mathcal{A}(x_1) \to (\mathcal{A}(x_1'))) \to (\forall x_1)\mathcal{A}(x_1))$.

for each *wf.* $\mathcal{A}(x_1)$ in which x_1 occurs free.

***Remarks* 5.3.1**

(*a*) The reader who is familiar with Peano's Postulates will recognise (N1), (N2) and (N7). Peano's Postulates are a set of axioms for the system of natural numbers which were first made explicit well before formal systems were studied as such. They are :

1. 0 is a natural number.
2. For each natural number n, there is another natural number n'.
3. For no natural number n is n' equal to 0.
4. For any natural numbers m and n, if $m' = n'$ then $m = n$.
5. For any set A of natural numbers containing 0, if $n' \in A$ whenever $n \in A$, then [illegible] contains every natural number.

Note that the first two postulates do not correspond with any of the axioms for our system N. We do not ned them because we have included symbols in the language $\mathcal{L}_N$ (0 and ', or a_1 and f_1^1) which must have interpretations in any model, so that in any model, an element $\bar{a}_1$ exists, and for each x there must be an element $f_1^1(x)$.

(*b*) The correspondence between (*N7*) and the fifth of Peano's postulates is not exact. Both are versions of the Principle of Mathematical Induction. However, because in *N* we are restricted to the use of the first oder language $\mathcal{L}_N$ the axiom (*N7*) cannot be as strong or inclusive as Peano's fifth postulate. The reason is that Peano's fifth postulate contains a second order quantifier 'for any set *A* of natural numbers', which cannot be expressed in our first order language. The best we can do is use the notion of axiom *scheme*, so that we effectively have a quantifier in 'for every *wf.* $\mathcal{A}(x_1)$ in which x_1 occurs free'. Note that such a *wf.* $\mathcal{A}(x_1)$ determines a set in any interpretation, namely the set of all elements v_1 of the domain of the interpretation which satisfy $\mathcal{A}(x_1)$. (More precisely, the set of all elements v_1 of the domain of the interpretation such that every valuation v for which $v(x_1) = v_1$ satisfies $\mathcal{A}(x_1)$.)

If we think in the context of a model of $\mathcal{N}$, therefore, each instance of the axiom scheme (*N7*) corresponds to the assertion of Peano's fifth postulate in regard to one particular set. However, there is still an essential difference. The instances of axiom scheme (*N7*) form a countable set of *wfs.* of $\mathcal{L}_N$. Peano's fifth postulate is a statement about all sets of natural numbers, and the collection of all these is uncountable. Thus (*N7*) is a much restricted form of the Induction Principle, since it refers only to that countable collection of subsets of the domain of a model which can be 'represented' in the manner described above by *wfs.* of $\mathcal{L}_N$.

(*c*) Peano's Postulates contain no mention of sums or products. These functions can be defined in terms of the successor function, using the induction principle, but it is convenient to include symbols for these in the formal language. Having done this, axioms (*N3*)-(*N6*) are necessary to ensure that in any model the interpretations of these symbols have the required properties.

▷ There is a fundamental difference mathematically between this situation and the situation with groups. The formal system of group theory allowed many different normal models, namely all groups. The system $\mathcal{N}$ of arithmetic is intended to have only one normal model, namely the set of natural numbers, since it is properties of natural numbers which we hope will appear as theorems in the system. Whereas the group theorist may be concerned with general results which hold in all groups, the number theorist is concerned with results about a particular set, the set of natural numbers. It is a natural question to ask, therefore, whether there are any normal models of the system $\mathcal{N}$ other than the set of natural numbers. Another question which arises naturally is whether the system is strong enough, in the sense of having as theorems all *wfs.* which we would like to be theorems, i.e. all *wfs.* which correspond to true statements about natural numbers. These two questions are not unconnected, as we shall see shortly.

Some readers may be familiar with the standard proof that Peano's Postulates determine the set of natural numbers uniquely. Let *N* and *M* be *'models' of Peano's Postulates. Then* $0 \in N$ and $0 \in M$. Let *A* be the set of elements of *N* which are elements of *M*. Then $0 \in A$. Also if $n \in A$, then $n \in N$ and $n \in M$, so $n' \in N$ and $n' \in M$, so $n' \in A$. Thus, by Peano's fifth postulate. *A* consists of all natural numbers, i.e. $A = N$, and so $N \subseteq M$. Similarly, $M \subseteq N$, and so $M = N$. In this proof, the fifth postulate is used essentially, and as we have noted above, (*N7*) does not correspond exactly to this postulate. Indeed the above proof cannot be translated into a proof in *N*. So there is no hope here of obtaining a negative answer to our first question about *N*.

Let us now address ourselves to the question : Is $\mathcal{N}$ complete? i.e. is $\mathcal{A}$ or $(\sim\mathcal{A})$ always a theorem of $\mathcal{N}$, for each closed *wf.* $\mathcal{A}$ of $\mathcal{L}_N$. The significance of this question may not be obvious at first sight, but it has a bearing on both questions above. If *N* were not complete, then it would not be a strong enough system in the above sense, for then there would be a closed *wf.* $\mathcal{A}$ such that neither $\mathcal{A}$ nor $(\sim\mathcal{A})$ were theorems of *N*. Now in any interpretation a closed *wf.* is either true or false, so in the interpretation *N* either $\mathcal{A}$ is true or $\mathcal{A}$ is false, and in the latter case $(\sim\mathcal{A})$ is true. Now the interpretation of $\mathcal{A}$ in $\mathcal{N}$ is a statement about natural numbers, and, intuitively, either $\mathcal{A}$ or $(\sim\mathcal{A})$ will have an interpretation which is a true statement about natural numbers. But neither $\mathcal{A}$ nor $(\sim\mathcal{A})$ is a theorem of $\mathcal{N}$. Thus if $\mathcal{N}$ were not complete then there would be a true statement about numbers whose corresponding *wf.* in *N* was not a theorem of $\mathcal{N}$. It would be desirable, and was part of the original aim in constructing the system *N*, that all the *wfs.* which are true in the model *N* should be theorem of *N*. However, if $\mathcal{N}$ were not complete then this could not be so.

Also, if there were a *wf.* $\mathcal{A}$ such that neither $\mathcal{A}$ nor $(\sim\mathcal{A})$ were theorems of *N*, then (provided that *N* itself is consistent) we could obtain, two distinct consistent extensions of *N* by adding first $\mathcal{A}$ as a new axiom and second $(\sim\mathcal{A})$ as a new axiom. Each of these extensions will have a normal model (Proposition 5.6), and these models are certainly models of *N* which must be essentially different, since in one $\mathcal{A}$ is true and in the other $(\sim\mathcal{A})$ is true. Thus if *N* were not complete there would necessarily be a normal model of *N* other than the intended one.

That *N* is not complete was one of the major results first obtained by Godel. In fact he proved a much stronger result with this as a special case.

5.4 Formal set theory

The foundations of mathematics are nowadays laid in the theory of sets, and since the beginning of this century mathematicians have investigated the basic assumptions that have to be made about sets (i.e. axioms) and the ways in which all of mathematics can be built upon these assumptions. The advantage of developing a formal theory of sets lies in making the assumptions explicit and providing an opportunity to criticise them andto explore interdependences between them. We shall describe one system of formal set theory. There are others, but ours is one of the standard ones, and it is perhaps easiest to describe, in terms of concepts we have already discussed. The reader who is unfamiliar with the set theoretic foundations of mathematics may find the axioms themselves difficult, but they are included here for the sake of completeness and in order to give some idea of their nature. What comes later does not depend on them. We do not have space to do more than describe the system and to point out some of the ways in which set theory develops from it.

The system which we describe is called *ZF*. The name derives from Ernst Zermelo, who first formulated a collection of axioms for set theory in 1905, and Abraham Fraenkel, who modified them in 1920.

The first order language which is appropriate for *ZF* contains variables, punctuation, connectives and quantifier as usual, and the predicate symbols = and A_2^2 (no function letters or individual constants). A_2^2 is intended to be interpreted as $\in$, the relation of membership. Indeed, with the same warning as was given in the case of $\propto_N$, we shall consider $\in$ as a symbol of the language, standin for A_2^2, and write $t_1\, t_2$ in place of $A_2^2(t_1, t_2)$, for any terms t_1 and t_2. Notice that the lack of individual constants and function letters means that the only terms are the variables, and the only atomic formulas are of the form $x_i = x_j$ or $x_i \in x_j$. This may seem excessively restricting, but the axioms which we introduce will ensure that the formal system genvinely reflects the full generality of intuitive set theory, and we shall be able to introduce defined symbols corresponding to the standard notions of set theory, such as the empty set, union, power set, etc.

ZF is defined to be the extension of $K_\propto$ (where $\propto$ is as described above) obtained by including as axioms (*E*7), all appropriate instances of (*E*9) (*E*8 ha no non-trival instances), and (*ZF*1) to (*ZF*8) listed below.

(*ZF*1) $(x_1 = x_2 \leftrightarrow (\forall x_3)(x_3 \in x_1 \leftrightarrow x_3 \in x_1))$.

This is called the Axiom of Extensionality, and its intended meaning is that two sets are equal if and only if they have the same elements. Note that the left to right implication is given already by (*E*9), but it makes the significance of this axiom clearer if we include both implications here.

(*ZF*2) $(\exists x_1)(\forall x_2) \sim (x_2 \in x_1)$.

This is the Null Set Axiom, sinc eit guarantees the existence, in the intended interpertation, of a set with no members. It is a consequence of (*ZF*1) that in any normal model there will be only one such set. We can thus introduce into the language the symbol $\varnothing$, to act as an individual constant, the *wf.* $(\forall x_2) \sim (x_2 \in \varnothing)$ being then the form that (*ZF*2) takes.

Notation, We introduce the symbol $\subseteq$ as an abbreviation as follows :

$(t_1 \subseteq t_2)$ stands for $(\forall x_1)(x_1 \in t_1 \rightarrow x_1 \in t_2)$

where t_1 and t_2 are any terms.

(*ZF*3) $(\forall x_1)(\forall x_2)(\exists x_3)(\forall x_4)(x_4 \in x_3 \leftrightarrow (x_4 = x_1 \vee x_4 = x_2)$

This is the Axiom of Pairing. Given any sets *x* and *y* there is a set *z* whose members are *x* and *y*. Again this axiom asserts existence, and it is convenient to introduce the symbols { and } into the language in order to denote the object whose existence is being asserted. $\{x_1, x_2\} \leftrightarrow (x_4 = x_1 \vee x_4 = x_2)$.

(*ZF*4) $(\forall x_1)(\exists x_2)(\forall x_3)(x_3 \in x_2 \leftrightarrow (\exists x_4)(x_4 \in x_1 \wedge x_3 \in x_4)$

This is the Axiom of Unions. Given any set x, there is a set y which has as its members all members of members of x.

Notation. We denote by $\cup x_1$ the object whose existence is asserted in (*ZF*4). This acts as a term, so $\cup$ acts as a one-place function symbol. We can then introduce $\cup$ by :

$(t_1 \cup t_2)$ stands for $\cup(t_1, t_2)$.

(*ZF*5) $(\forall x_1)(\exists x_2)(\forall x_3)(x_3 \in x_2 \leftrightarrow x_3 \subseteq x_1)$

This is the Power Set Axiom. Given any set x there is a set y which has as its members all the subsets of x.

(*ZF*6) $(\forall x_1)(\exists x_2)\,\mathcal{A}(x_1, x_2) \rightarrow$

$$(\forall x_3)(\exists x_4)(\forall x_5)(x_5 \in x_4 \leftrightarrow (\exists x_6)(x_6 \in x_3 \wedge \mathcal{A}(x_6, x_5))),$$

for every *wf.* $\mathcal{A}(x_1, x_2)$ in which x_1 and x_2 occur free (and in which, we may suppose without loss of generality, the quantifiers $(\forall x_5)$ and $(\forall x_6)$ do not appear).

This is the Axiom Scheme of Replacement. If the *wf.* $\mathcal{A}$ determines a function, then for any set x, there is a set y which has as its members all the images of members of x under this function.

(*ZF*7) $(\exists x_1)(\varnothing \in x_1 \wedge (\forall x_2)(x_2 \in x_1 \rightarrow x_2 \cup \{x_2\} \in x_1))$.

(N.B. $\{x_2\}$ is an abbreviation of $\{x_2, x_2\}$, defined above.)

This is the Axiom of Infinity. It asserts the existence, in any model, of an infinite set. If it were not included amongst the axioms there would be no way of ensuring that the formal system had any relevance to intuitive set theory which includes infinite sets.

(*ZF*8) $(\forall x_1)(\sim x_1 = \varnothing \rightarrow (\exists x_2)(x_2 \in x_1 \wedge \sim (\exists x_3)(x_3 \in x_2 \wedge x_3 \in x_1)))$.

This is the Axiom of Foundation. Every non-empty set x contains a member which is disjoint from x. This is a technical axiom which is included in order to a avoid anti-intuitive anomalies such as the possibility of a set being a member of itself.

ZF is a formal system of set theory. The axioms are chosen so that the interpretations of the formal symbols in normal models will behave as sets do. Some of the axioms have a stronger basis in intuition than others, but these eight have stood the test of time as representing basic truths about sets.

ZF can be used as a basis for mathematical analysis in the following way. On the assumption that it is a consistent system, we know that a normal model exists. It can be shown that in any such model there are sets with all the usual properties of the number systems. The details of this are lengthy and cannot be covered here. For example, a model for the system $\mathcal{N}$ of arithmetic can be defined as a subset of a model of *ZF* in the following way. $\varnothing$ has an interpretation in the model of *ZF*, $\overline{\varnothing}$, say. Then $\{\overline{\varnothing}\}$ is a different element of the model (the set whose only member is $\overline{\varnothing}$), $\{\overline{\varnothing}, \{\overline{\varnothing}\}\}$ is another (this set has two elements, $\overline{\varnothing}$ and, $\{\overline{\varnothing}\}$). This is the beginning of an inductive process, generating a sequence of sets. The general rule is : for each x in the sequence, its successor is $x \cup \{x\}$. It can be verified easily that the $(k+1)$st member of this sequence has k elements, and it is possible to *define* the natural number k as this $(k+1)$st member. We have already seen that the other arithmetic operations can be defined in terms of the successor function. The axioms (*N*1).... (*N*7) are then consequences of the definitions and the *ZF* axioms. Note that (*ZF*7) is needed to ensure that the collection of all members of this sequence is an element of our normal model of *ZF*. In this way every normal model of *ZF* contains a normal model of $\mathcal{N}$.

The reader with a mathematical background may be familiar with the way in which the number systems of integers, rationals, reals and complex numbers may be constructed, starting from the natural members, by algebraic procedures. All of these procedures can be carried out within *ZF*. There is a lot of detailed verification required but the end result is confirmation that every normal model of *ZF* contains a set which looks and behaves like the set of complex numbers. (This set of course has a subset which looks and behaves like the set of real numbers.)

Besides the foundation of analysis on an axiomatic base, there was another stimulus at the turn of the century for the study of axiomatic set theory and this was in the intuitive justification (if any) for the use of certain principles in mathematics. Attention was then focussed on two particular principles, the axiom of choice (which was known to have several equivalent formulations) and the continuum hypothesis. It is quite illuminating to investigate some of the history of these principles since that time. Some mathematicians have regarded them as additional axioms of set theory, and others have regarded them either as suspect intuitively or even as falsehoods.

(*AC*) For any non-empty set x there is a set y which has precisely one element in common with each member of x.

(Two of the best known equivalent formulations are : *Zorn's Lemma;* if each chain in a partially ordered set has an upper bound, then there is a maximal element of the set, and *The Well-Ordering Principle;* each set can be well-ordered.)

The continuum hypothesis is :

(*CH*) Each infinite set of real numbers either is countable or has the same cardinal number as the set of all real numbers. (Two sets have the same cardinal number if there is a bijection between them.)

Because mathematicians were not in agreement about the acceptability of these two principles, the question of course was asked; are they true? The next question is: if these principles are to be demonstrated, on what principles ought such demonstrations to be based? Zermelo and Fraenkel (and others) listed what they thought to be the fundamental principles of set theory, and the question became; Can (*AC*) and (*CH*) be deduced as theorems of the system *ZF* of set theory, and if they can not, would it be consistent to include one or both as additional axioms?

5.5 Consistency and models

Any first order system is consistent if and only if it has a model. It is possible to argue, then, that the mathematical systems we have described are consistent because in each case we have been mirroring, in the axioms, the properties of an intended model. However, the perceptive reader may already have been concerned about a possible circularity in our arguments, which can be exemplified by the definition in Unit 3 of an interpretation as a *set* with certain operations and relations. How can we talk of interpretations or models of *ZF*, the formal system of set theory, then, without a circularity? The answer is in the ideas previously mentioned of a metatheory embodying the assumptions which have to be made in order to prove results about formal systems. When we deal with the system *N*, for example, it is possible to use, say, *ZF* as a metatheory since *ZF* 'cotains' *N* in a sense already referred to. However, when we discuss *ZF* we have, so to speak, reached the end of the line. By its nature, *ZF* is to be appropriate to set theory and hence for all of mathematics. However, in order to study *ZF* we require mathematical methods which are not part of *ZF*. The notion of an interpretation of *ZF* can be defined only in terms of some intuitive metatheory concerning 'real' sets. The elements of a model of *ZF* are to be thought of as sets interpreting the symbols of *ZF*. However, the domain of a model of *ZF*, though it may be a 'real' set, cannot be a set in the same sense that the elements of that domain are, for it cannot be the interpretation of a symbol of *ZF*.

There are certainly intuitive and semantic difficulties in these matters, and it is because of this that demonstrations of consistency by means of models are generally held to be inadequate. The more respectable approach is the following. Given two first order systems *S* and S^*, we may appempt to show, on the assumption that a model exists for S^*, that a model for *S* can be constructed. This would give a proof of *relative consistency*. There is one situation where this is almost trivial.

***Proposition* 5.5.1**

Let S^* be an extension of *S*. Then if S^* is consistent, so is *S*.

Proof. Suppose that S^* is consistent, but *S* is not. Then $\vdash_S \mathcal{A}$ and $\vdash_S (\sim\mathcal{A})$, for some *wf.* $\mathcal{A}$ of *S*. But $\mathcal{A}$ is a *wf.* of S^* also, and any proof in *S* is also a proof in S^*, so $\vdash_{S^*} \mathcal{A}$ and $\vdash_{S^*} (\sim\mathcal{A})$ contradicting consistency of S^*.

This is the easiest situation to deal with. In cases where S^* is not an extension of *S* in this sense, for example, where the languages of the two systems are different, the proof of relative consistency would be more difficult and may involve actual construction of a model of *S* from an assumed model of S^*. We have given one such construction, albeit sketchily, in showing that consistency of *ZF* implies consistency of *N*.

It is not known whether *ZF* is consistent. Most logicians believe that it is, but any attempt to prove that it is consistent will lead to difficulties of the kind described above. Essentially, it would require the asumption of consistency of a system even more all embracing than *ZF*. Certainly there would be no corresponding difficulties in the way of an attempt to disprove consistency. All that would be required for that would be an example of a *wf.* $\mathcal{A}$ such that both $\mathcal{A}$ and $(\sim\mathcal{A})$ are theorems of *ZF*. It is implicit in the above that no such *wf.* has yet been found. Seventy years of fruitless search is evidence that no such *wf.* exists, but it is no way conclusive.

Finally let us note a result about models of *ZF*. *ZF* is a first order system. Under the assumption that *ZF* is consistent, *ZF* has a *countable* model. Now uncountable sets exist, intuitively, so we would expect models of *ZF* to be uncountable, in order to contain such sets. This apparent paradox is called Skolem's Paradox, but we can escape from a direct contradiction by careful consideration of what constitutes a model, in the following way.

To be specific, axiom (*ZF*5) is interpreted as 'given any set *x*, there is a set consisting of all subsets of *x*'. If *x* is an infinite countable set, then, according to the rules of set ehtory, *x* has uncountably many subsets. How can the set of all subsets of such a set *x* belong to a countable model? A countable model of *ZF* consists of sets. For any 'real' set *x* which belongs to the model (clearly) there must be 'real' sets which do not belong to the model), axiom (*ZF*5) asserts that all of the subsets of *x* which belong to the model constitute a set *y* which must also belong to the model. This set *y* will be countable, when regarded as a 'real' set, but it will be uncountable when regarded as an element of the model. An infinite set is uncountable if there is no bijection between it and the set of natural numbers. In the model there will be no bijection between *y* and the set of natural numbers (all 'real' such bijections will be missing from the model in the same way that some subsets of *x* are missing).

Summary :

- It *S* is a consistent first order system with equality, then *S* has a model in which the interpretation of A_1^2 is =
- Let *S* be a first order system with equality. Then the following are theorems of *S*.

  (i) $(\forall x_1)A_1^2(x_1, x_2)$

  (ii) $(\forall x_1)(\forall x_2)(A_1^2(x_1, x_2) \rightarrow A_1^2(x_2, x_1))$,

  (iii) $(\forall x_1)(\forall x_2)(\forall x_3)(A_1^2(x_1, x_2) \rightarrow A_1^2(x_2, x_3) \rightarrow A_1^2(x_1, x_3)))$.
- Let *S* be a first order system with equality. A normal model of *S* is a model in which A_1^2 is interpreted as =.
- Let S^* be an extension of *S*. Then if S^* is consistent, so is *S*.

●●●