PART I

This part serves two purposes. It presents some basic material on truth-functional logic, and it also makes clear our stand on various general issues in the philosophy and methodology of logic. The basic material includes: the description of a language for truth-functional logic; definitions of proof and validity and of derivation and consequence for such a language; and proofs of completeness, decidability and compactness. The general material deals with the aim of logic, the nature of such central concepts as abbreviation, deduction and consequence, and the relation of truth-functional logic to other systems.

Chapter 1. Syntax

We shall be interested in two central aspects of language: grammaticality and meaning. Two important questions concerning grammaticality are: what are the grammatical expressions of a given language; and what is their grammatical analysis? Similarly, two important questions concerning meaning are: what are the meanings of the grammatical expressions; and how, if at all, are these meanings to be determined on the basis of the grammatical analysis of the expressions that have them? Syntax deals with questions of the first sort, semantics with those of the second. It is common, both in logic and in linguistics, to require that syntax be prior to semantics in the sense that the grammatical expressions are to be determined without regard to their meaning or, indeed, to any feature but their form. We largely adopt this standard view, though we do consider an alternative approach in the exercises.

The syntax of a language has two parts: the first is a characterization, usually in the form of a list, of its symbols - or alphabet; and the second is a set of rules for determining which strings of symbols are grammatical. In logic, these rules are called formation rules and are said to determine the formulas. Let us describe these two parts of the language **L** for truth-functional logic.

The alphabet of **L** consists of the following symbols:

(i) the letters (or sentence letters) $\mathbf{p_1}$, $\mathbf{p_2}$,...,

(ii) the truth-functional connectives $\vee$ and $\neg$,

(iii) the left and right hand brackets **(** and **)**.

The formation rules of **L** are given in the following inductive definition. (A general account of inductive definition is given in the appendix).

definition *a.* (i) Each sentence letter is a formula;

(ii) If **A** is a formula so is $\neg$**A**;

(iii) If **A** and **B** are formulas then so is **(A $\vee$ B)**.

Thus formulas are built up from sentence letters by placing $\neg$ before a formula and by inserting $\vee$ between two formulas and enclosing the result in brackets. As noted in the appendix, we can show that an object satisfies an inductive definition by exhibiting a record of its generation. In the case of formulas such records can be regarded as grammatical analyses of the formulas generated. For example, the following record shows that **¬(**$\mathbf{p_1}$**∨¬**$\mathbf{p_3}$**)** is a formula and thereby provides a grammatical analysis of it.

1. p_1		(clause i)
2. p_3		(clause i)
3. $\neg p_3$	2	(clause ii)
4. $(p_1 \vee \neg p_3)$	2,3	(clause iii)
5. $\neg(p_1 \vee \neg p_3)$	4	(clause ii)

On the other hand, it should be clear that neither ∨ nor **((**$\mathbf{p_1}$ is a formula. We refer to non-empty strings of symbols from the alphabet of a language (whether grammatical or not) as that language's *expressions*.

We adopt some standard terminology for formulas. ¬ is a unary connective: it attaches to a formula to produce a formula. ∨ is a binary connective: it combines with two formulas to produce a formula. **¬A** is a negation with negated part **A**. **(A∨B)** is a disjunction with left disjunct **A** and right disjunct **B**.

The language **L** that we are talking about is said to be the object language. The language in which we talk about the object language is said to be the metalanguage. We follow the convention that we may refer in the metalanguage to any symbolic expression of the object language by printing it in boldface. Thus, the sentence letter $\mathbf{p_4}$ and the formula **(**$\mathbf{p_1}$ **∨** $\mathbf{p_2}$**)** are the expressions shown on the line below.

p_4 $(p_1 \vee p_2)$

This conforms with our general practice of making all the special symbols of the metalanguage boldface. Occasionally (when we want to talk about expressions that are themselves boldface, for example), we shall also follow the convention that an expression may be named by enclosing it in single quotes. (The latter convention may be more convenient for students should they wish to distinguish between formulas and formula-names in their own writing.)

We have said that the alphabet of our object language is comprised of "symbols" and that its formulas are "strings" built up by the appropriate placement of symbols. There are several ways in which one might attempt to make the notions of symbol and placement more precise. One approach is outlined in the exercises.

The letters '**A**' and '**B**' used in definition 1.a are called metalinguistic variables. They are not themselves particular expressions from the object language, nor even the names of such expressions. Rather they are variable-signs from the metalanguage, used in the statement of formation rules to range over arbitrary expressions from the object-language, and used thereafter in this chapter to range over arbitrary formulas from the object language. So, for example, the assertion that **(A ∨ A)** is a disjunction with left disjunct **A** does not imply that the symbol '**A**' is a formula. It does imply, however, that **(**$\mathbf{p_1}$ **∨** $\mathbf{p_1}$**)** is a formula with left disjunct $\mathbf{p_1}$, that **((¬**$\mathbf{p_3}$ **∨ ¬**$\mathbf{p_1}$**) ∨ (¬**$\mathbf{p_3}$ **∨ ¬**$\mathbf{p_1}$**)** is a formula with left disjunct **(¬**$\mathbf{p_3}$**∨¬**$\mathbf{p_1}$**)**, and similarly for all the other formulas. We also take $\mathbf{p_i}$ to indicate the i-th sentence letter, i.e, **p** subscripted by the arabic numeral for i. (Note that it is the numeral rather than the number that subscripts **p**.) Similar conventions will be adopted without notice elsewhere.

The immediate components of a formula are those formulas, if any, from which it is directly constructed. Thus the immediate components of $\mathbf{(A \vee B)}$ are **A** and **B**, and the sole immediate component of $\mathbf{\neg A}$ is **A**. $\mathbf{p_4}$ has no immediate components because it is not constructed out of anything, but is rather a primitive symbol. The subformulas of a formula **A** are all the formulas involved in its construction. Thus **A**'s subformulas are **A** itself, **A**'s immediate components, the immediate components of those components, and so on. For example, the subformulas of $\mathbf{((\neg p_1 \vee p_2) \vee p_1)}$ are $\mathbf{p_1}$, $\mathbf{\neg p_1}$, $\mathbf{p_2}$, $\mathbf{(\neg p_1 \vee p_2)}$, $\mathbf{((\neg p_1 \vee p_2) \vee p_1)}$, but not $\mathbf{(p_2 \vee p_1)}$.

For any formula **A**, **SUB(A)** is to be the set of all subformulas of **A**. If $\mathbf{\Gamma}$ is a set of formulas, $\mathbf{SUB(\Gamma)}$ will be the set of all subformulas of the members of $\mathbf{\Gamma}$. The truth-functional constituents of **A** are the subformulas of **A** that are neither disjunctions nor negations. (The word **truth-functional** will be omitted when the intended qualification is clear from context.) For example $\mathbf{((\neg p_1 \vee p_2) \vee p_1)}$ has two truth functional constituents, $\mathbf{p_1}$ and $\mathbf{p_2}$. **CST**(**A**) is to be the set of constituents of **A**, and **CST**(Γ) the set of constituents of members of $\mathbf{\Gamma}$. In **L**, the constituents of a formula are its sentence-letters, but in chapter 8 we will consider languages in which this is not the case. It is often useful to distinguish between expressions and their occurrences. For example, although $\mathbf{((\neg p_1 \vee p_2) \vee p_1)}$ has only two constituents, viz. $\mathbf{p_1}$ and $\mathbf{p_2}$, it has three occurrences of constituents, since $\mathbf{p_2}$ occurs once and $\mathbf{p_1}$ occurs twice. We use underlining to indicate occurrences. Thus $\underline{\mathbf{A}}$ will indicate an occurrence of **A**. Our philosophy is to take talk of occurrences as basic. However, it is of interest to see if one can dispense with reference to occurrences; and a method of so doing is outlined in the exercises.

We turn to the concept of scope. Intuitively, the scope of a symbol occurrence in an expression is the portion of the expression that is "governed" by the symbol occurrence (we may include the symbol occurrence itself). In the case of the language **L**, the scope of an occurrence of a connective may be defined to be the smallest subformula occurrence to contain the connective occurrence. For example, in the formula $\mathbf{(\neg(p_1 \vee p_3) \vee \neg p_2)}$, the scope of the leftmost disjunction is the occurrence of $\mathbf{(p_1 \vee p_3)}$ and the scope of the leftmost negation sign is the occurrence of $\mathbf{\neg(p_1 \vee p_3)}$. If the scope of one connective occurrence is contained within that of another the second is said to have wider scope than the first. We also say in this case that the two connective occurrences are nested. So in $\mathbf{(\neg(p_1 \vee p_3) \vee \neg p_2)}$ the two disjunctions are nested, but the two negations are not. Finally, in every formula there is one connective occurrence whose scope is the entire formula occurrence. If the formula is a disjunction this connective will be $\vee$ and if it is a negation it will be $\neg$. We sometimes call this connective the main connective of the formula.

If we take a formula **A** and replace one or more of its sentence letters by formulas in such a way that each sentence letter is replaced by the same formula at every occurrence we obtain a substitution instance of **A**. For example **((¬¬p_1∨¬p_2)∨¬p_1)** is a substitution instance of **((¬p_1∨p_3)∨p_1)**. (¬p_1 is substituted for p_1 and ¬p_2 is substituted for p_3.) It is not a substitution instance of **((¬p_1∨p_1)∨p_1)**, nor of **((¬p_1∨¬¬p_3)∨p_1)**. Similarly, if **Γ** is a set of formulas, a substitution instance of **Γ** is a set obtained by uniformly substituting formulas for sentence-letters (i.e., substituting so that a sentence letter is replaced by the same formula at every occurrence) in all the members of **Γ**. For example **{¬(p_1∨p_3),(¬(p_1∨p_3)∨p_2)}** is a substitution instance of **{¬p_1,¬(p_1∨p_2)}**; **{¬p_1,¬(p_3∨p_2)}** is not. A pair **Γ',A'** is a substitution instance of the pair **Γ,A** if **Γ'** and **A'** result from a single uniform substitution of formulas for sentence-letters in **Γ** and **A**, respectively.

Let us say that an occurrence of an expression **E** in a formula **A** is a syntactic occurrence if it is involved in a construction of **A** according to the rules (i) - (iii) of definition . An occurrence that is not syntactic may be termed merely typographic. For example, the occurrence of **¬(p_1∨** in **¬(p_1∨p_2)** is merely typographic whereas the occurrence of p_1 is syntactic. We then have the following significant result.

Theorem b. (Transparency). Any occurrence of a formula within a formula is a syntactic occurrence.

We leave the precise definition of syntactic occurrence and the proof of this result as a problem. One important consequence of the result (not to be confused with the result itself) is that the formulas of **L** are unambiguous; there is only one way they can be constructed, or read:

Corollary c. (Unique Readability). For each formula **A** exactly one of the following holds: it is a sentence-letter; it is of the form **¬B** for unique **B**; it is of the form **(B∨C)** for a unique pair of formulas **B** and **C**.

Proof. The only problematic case is when **A** is of the form **(B∨C)** and of the form **(D∨E)** for distinct pairs of formulas **B**,**C** and **D**,**E**. If this is so, then either **B** will properly begin **D** or **D** will properly begin **B**. Without loss of generality, suppose the former. By the transparency theorem, **B** occurs in a construction of the formula **D** = **D**(**B**). So replacing **B** with the letter **p**, it follows that **(D**(**p**), which is of the form **pF** for some expression **F**, is a formula. But it is readily shown that **p** is the only formula that begins with **p**; and so **D** = **D**(**B**) is the same as **B**, contrary to supposition.

Two other methods of proving this result, not using transparency, are outlined in the problems.

Another important consequence of the transparency theorem is that the result of replacing the occurrence of a given formula in a formula by some other formula is itself a formula. This consequence was implicitly taken for granted in our discussion of replacement and substitution above. But if the occurrence of the formula to be replaced is merely typographic, then it is not evident that the result of the replacement will also be a; and so again, something like Transparency is required to guarantee that the result will hold.

Drills, exercises, and problems

1**[d]**. Give examples of syntactic and semantic remarks that are true of English.

2**[d]**. Determine which of the expressions exhibited below is a formula of **L** and provide grammatical analyses for those that are.

a. $)p_1 \vee p_2($
b. $(p_1 \vee (\neg p_8 \vee p_6))$
c. (p_{107})
d. $(((\neg\neg\neg p_1 \vee p_1) \vee p_1) \vee p_1)$
e. $\neg(p_6 \vee p_9) \vee (p_9 \vee p_6)$
f. $\neg\neg\neg(p_3 \vee p_1)$
g. $(A \vee B)$

3**[d]**. List the immediate components and the subformulas of the following formulas:

a. $\neg(p_1 \vee p_6)$
b. $((p_2 \vee \neg\neg\neg p_2) \vee p_3)$
c. p_9

4**[e]**. Identify the symbols of **L** that can occur first in a formula. Prove by induction that no other symbols can.

*5**[e]**. Prove rigorously that $\mathbf{(p \vee q \vee r)}$ is not a formula. [Hint: Find a property that all formulas have and $\mathbf{(p \vee q \vee r)}$ lacks. Use induction to prove that all formulas have the property.]

6**[d]**. State whether the claims below are true or false and give reasons, in each case, for your answer:

a. $\mathbf{(A \vee \neg B)}$ is a disjunction.
b. The result of prefixing 'p_1' with two dashes is a formula.
c. Every formula has at least one subformula.

d. **A** is a subformula of **(A∨B)**.

e. If **A** is a subformula of **B** and **B** is a subformula of **A** then **A** and **B** are identical.

f. Every substitution instance of a disjunction is a disjunction.

g. If **A** is a substitution instance of **B** and **B** is a substitution instance of **C** then **A** is a substitution instance of **C**.

h. If one substitution instance of a formula is a disjunction, then every substitution instance is.

i. If two formulas have the same substitution instances, they are the same.

j. If **A** ε **Γ** then **SUB(A)** is a subset of **SUB(Γ)**.

7**[d]**. List all pairs of the formulas below which are such that the first is a substitution instance of the second.

a. p_2

b. $(p_1 \vee p_2)$

c. $(p_1 \vee p_1)$

d. $(p_2 \vee p_1)$

e. p_1

*8**[e]**. Show that there are infinitely many formulas, no two of which have a common substitution instance. [Hint: let $\mathbf{A_1}$,$\mathbf{A_2}$,... be a sequence of formulas of increasing length whose sole sentence letter is **p**. Consider the formulas $\mathbf{(p \vee A_1)}$, $\mathbf{(p \vee A_2)}$, ...]

9**[e]**. A strict rendering of clause (ii) in the definition of a formula would be:

if **A** is a formula then so is the result of juxtaposing ¬ and **A** in the order indicated.

What would be a strict rendering of clause (iii)?

10**[e]**. Give an example of a formula with three occurrences of disjunction, such that the leftmost occurrence has wider scope than the middle and the middle has wider scope than the third.

11**[e]**. A <u>substitution function</u> is defined to be a function from sentence letters to formulas.

(i) Given a substitution function S, define the natural extension of S to a function S^* on arbitrary formulas whose sentence letters lie in the domain of S.

(ii) Given a formula **A** and a substitution instance **B** of **A**, show that there is exactly one substitution function S whose domain consists of the sentence-letters of **A** and for which S(**A**) = **B**.

(iii) Hence show how to define the notion of substitution instance in terms of the notion of substitution function.

*12**[e]**. A language that allows outermost parentheses to be omitted can be defined as follows:

Every sentence letter is an "unmodified" formula.

If **A** is an unmodified formula then ¬**A** is an unmodified formula.

If **A** and **B** are unmodified formulas then so is **(A∨B)**, and if **A** and **B** are unmodified formulas then **A∨B** is a modified formula.

a) Show that transparency (1.b) fails relative to the grammar for this language. b)Explain why our definition of scope is no longer correct. How should it be modified? d)Show that the result of replacing either a syntactic or merely typographic occurrence of a formula by a formula is still a formula.

*13**[e]**. (strings). By the juxtaposition of two expressions is meant the result of placing the one next to the other (in left to right order). Thus if one of the expressions is the string of symbols $\mathbf{E_1}...\mathbf{E_m}$ and the other is $\mathbf{F_1}...\mathbf{F_n}$, then their juxtaposition is the string $\mathbf{E_1}...\mathbf{E_m}\mathbf{F_1}...\mathbf{F_n}$. The word **the**, for example, is the juxtaposition of **t** and **he** and also the juxtaposition of **th** and **e**. Let us use the symbol '(**E***F**)' for the juxtaposition of **E** and **F** (dropping brackets, if need be, in the usual way). Take '**A**', '**B**', '**C**' as metalinguistic variables whose range is the expressions of **L** (i.e., the strings of symbols of the alphabet of **L**).

a. Declare the following conditions true or provide a counterexample: (i) If **A≠B** then **A*B≠B*A** (ii) If **(A*B)*C**=**A*(B'*C)** then **B**=**B'**.

b. Rewrite definition 1.a using the new notation. What convention makes it possible to do without asterisks in the text?

c. Prove that every expression of **L** has a unique representation of the form A_1*...*A_n, where A_1,...,A_n belong to the alphabet of **L** and the terms are associated to the left.

*14**[p]**. (the geometry of symbols and expressions).

It is possible to think of symbols and expressions in geometric terms. One natural idea is to identify each symbol with an appropriate two-dimensional shape. To this end, we first try to define a shape. We do not attempt to give a "philosophical" definition, according to which a shape is an abstraction from something more concrete, such as physical objects, because of the difficulty in making the concepts involved precise. Instead, as a first attempt, we take a shape to be a class of geometrically similar figures in two-dimensional Euclidean space. (The notions of "figure" and "similarity" are considered given.) This idea will not do in the case of symbols like **p** and **d**, which correspond to the same shape.

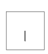
We might therefore think of a symbol as an oriented shape. a. How might an oriented shape be defined? (Hint: Take an oriented shape to be a similarity class of oriented figures, where an oriented figure is a special kind of figure, consisting of an unoriented figure enclosed in a square whose sides are labeled "left", "up", "right", and "down" in a clockwise manner.) Notice that even if our alphabet is such that orientation is not needed to identify symbols, it is still needed to identify expressions that are "left-right" juxtapositions of more than one symbol. b. How might the notion of shape be modified so that expressions of more than one symbol can be identified with shapes. (Hint: Say that two oriented figures are of the same scale if their squares are congruent. Let a "row" be a figure composed of a series of oriented figures of the same scale, the right side of each constituent oriented figure coinciding with the left side of its successor, if it has one. Let a shape now be a similarity class of rows. It is then easy to describe the shape to be identified with the expression formed by juxtaposing symbols $s_1,...,s_n$ in order.) c. In some languages there are "upper case" and "lower case" versions of the same symbols. To accommodate this feature we might simply stipulate that certain pairs of oriented figures are to be regarded as representatives of upper and lower case versions of particular symbols. Frequently (as with the Roman **s** and **S**) upper and lower case versions of a symbol are distinguished by their relative sizes. Notice that the square (that was introduced to provide figures with an orientation) allows us to distinguish similar shapes of different sizes. There is an element of idealization here, since in ordinary language there is no way to distinguish between occurrences of **s** and **S** in isolation. An equally significant idealization is that all "instances" of symbols and expressions are taken to be exactly similar. The problems in defining a symbol so as to include all its ordinary readable instances would be considerable.

15**[p]**. (design-a-symbol) The expressions of **L** are constructed from an alphabet of symbols, which, in the previous problem, are identified with oriented shapes. It is possible to regard the symbols of an alphabet as themselves constructed from a smaller set of "typographic atoms". For example **d** can be regarded as the "elision" of a **|** and a **o**, and **e** as a "merger" of a **c** and a **-**. Describe an economical set of typographic atoms and operations on them from which idealized versions of the digits 0-9 can be constructed. (Hint 1 ("tiling" approach): Let the typographic atoms be oriented shapes smaller than symbols, which we might call "tiles". In particular take the three tiles depicted below as atoms.

blank **arc** **line**

Let these atoms be the simplest rectangular "blocks" of tiles, and let other blocks be constructed by means of two operations. ROTATE turns a block 90^o in a clockwise direction. ADJOIN takes two blocks of the same height and forms a new block in which the right side of the first coincides with the left side of the second. Each digit can be regarded as a four-by-four block of sixteen tiles. For example, using superscripts to indicate the number of repeated applications of an operation and writing **[x_1+x_2+...+x_n]** for ADJOIN(...(ADJOIN(x_1,x_2),...,x_n)...), the symbol **nine**, depicted below, can be defined as: [**empty**+**highcircle**+**leftline**], where **empty** is ROTATE([**blank**+**blank**+**blank**+**blank**], **highcircle** is [ROTATE([ROTATE(**arc**)+**arc**+**blank**+**blank**])+ROTATE([ROTATE^2(**arc**)+ROTATE^3(**arc**)+**blank**+**blank**]) and **leftline** is ROTATE([**blank**+ROTATE(**line**)+ROTATE(**line**)+ROTATE(**line**)]).

Hint 2 ("overlay" approach): Let the typographic atoms be oriented shapes of the same size as the symbols themselves. In particular, take as atoms the two orientated shapes shown below.

line **arc**

Consider six operations on oriented shapes. LEFT, RIGHT, UP, and DOWN move the figure inside the orienting squares in the appropriate direction by an amount equal to one fourth of a side of the square. ROTATE turns the interior 90^o clockwise. MERGE "superimposes" two oriented shapes. For example, writing **[x_1*x_2*...*x_n]** for **MERGE(...(MERGE(x_1,x_2),...,x_n)...)**, **nine** may be defined [UP(**zero**)*RIGHT(**one**)] where **zero** is [[**arc***ROTATE(**arc**)]*[ROTATE^2([**arc***ROTATE(**arc**)])] and ONE is [**line***DOWN(**line**)*UP(**line**)*UP^2(**line**)].)

b. Using additional operations if necessary, show how arbitrary numerals and numerical subscripts can be constructed. (HINT: One approach would allow subscripted symbols to be bordered by oriented rectangles of arbitrary width. An alternative approach would retain the idea that the symbols of our alphabet are similarly sized, but would allow numerals within subscripts to be arbitrarily small.)

c. Using additional typographical atoms and operations if necessary, show how an idealized version of the alphabet of **L** can be constructed.

16**[p]**. (Abstract approach to syntax). We need take no particular view as to the nature of symbols or of expressions (and perhaps, from a purely logical standpoint, this is to be preferred). Suppose we specify only that **p_1**,**p_2**,... are distinct expressions (whatever they may be), that ∨ is a binary operation on expressions, and that ¬ is a unary operation on expressions. We may then inductively define the formulas along the lines of definition : each sentence letter is a formula; and if A and B are formulas then so are ¬(A) and ∨(A,B). a. What constraints must be placed on ∨, ¬, and the $\mathbf{p_i}$'s to ensure unique readability? b. Give an interpretation that fits these conditions in which formulas are natural numbers and ∨ and ¬ are numerical functions. (Such an interpretation is called a Gödel numbering.)

*17**[p]**. (naming conventions)

a. According to the boldface convention (which we have adopted in the text), the name of an expression may be formed by making it boldface. Explain why a boldface expression with more than one symbol is ambiguous and how the ambiguity may be resolved. According to the quotation mark convention, the name of an expression may be formed by enclosing it in quotation marks. How is the ambiguity noted above resolved under the quotation mark convention? Why is it easier to form names of names of expressions under the quotation mark convention?

b. We use boldface letters, such as **E** and **F**, as variables for expressions. Why does this conflict with the boldface convention. How should this convention be amended?

c. There are two very different ways of understanding our use of meta-linguistic variables. On one (the objectual approach), the variables take expressions as <u>values</u>. Suppose '**E**' takes the value **E** and '**F**' takes the value **F**. The expression '**EF**' may then be taken to denote the juxtaposition of **E** and **F** as long as we adopt the juxtaposition convention, according to which the juxtaposition of two denoting expressions denotes the result of juxtaposing their denotations. On the other approach (the substitutional), the meta-linguistic variables take boldface names of expressions as their <u>substituends</u>. Suppose '**E**' is assigned '**¬**' as substituend, for example, and '**F**' is assigned '**p**' as substituend. Then '**EF**' will be assigned the result of substituting '**¬**' for '**E**' and '**p**' for '**F**' in '**EF**', i.e. '**¬p**'. Thus there is no need for a juxtaposition convention in this case.

Suppose capital plainface letters are variables whose substituends are (plainface) expressions of a certain sort. Then what further convention governing boldface variables will ensure that capital boldface letters are variables whose substituends are boldface names of expressions of that sort? How can this further convention be elaborated so that **E**E' can have '**p**p' as an instance but not '**p**q'? Use this elaboration of the convention to make sense of the claim: if S is an English sentence, **S** is true iff S.

18**[p]**. (occurrences). We employ the following three locutions concerning the occurrences of expressions:

(i) (type) **e** is an occurrence of **E**;

(ii) (replacement) **F'** is the result of replacing the occurrence **e** in **F** with **E'** (**F'** may be taken to be **F** itself if **e** does not occur in **F**);

(iii) **g** is the result of adjoining the occurrences **e** and **f**. (Thus **g** is another occurrence and is only defined when **f** occurs immediately after **e**.)

a. In terms of these three basic notions alone, show how to define: **e** occurs in **F**; **e** occurs exactly twice in **F**; **e** precedes **f**; **e** and **f** overlap.

b. Suppose that **e** and **e'** are two occurrences of **E**. Then the respective suboccurrences of **e** and **e'** are aligned in an obvious way. Thus in **((p∨p) ∨ (p∨p))**, there are two occurrences of **(p∨p),** with the first and second occurrences of **p** in the first respectively aligned with the first and second occurrences of **p** in the second. Show how to define the alignment of occurrences in terms of our basic notions.

c. We present a way of representing occurrences. Suppose that **e** is an occurrence in **F**. Then **F** is of the form **GEH**, where **E** is the substring of **F** corresponding to the occurrence **e**. We identify **e** with the triple <**G**, **E**, **H**>. Show how to define the three basic notions above under this representation.

*19**[p]**. a) Suppose clause (iii) in the definition of formula of **L** had read: If **A** and **B** are formulas then so is **A** ∨ **B**. Show that unique readability would fail. b) Suppose that **p**, **p∨(p**, and **p)** were all sentence letters. Show that unique readability would again fail. c)Write a definition for formulas that permits one to omit outermost brackets (but not interior brackets). Show that the unique readability property continues to hold.

20**[p]**. (Transparency) a. Given an expression **E**, let nl(**E**) be the number of left-hand brackets in **E** and let nr(**E**) be the number of right-hand brackets in **E**. Say that the expression **E** <u>properly begins</u> a formula **A** if **A** consists of **E** followed by further symbols and that **E** properly ends **A** if **A** consists of **E** preceded by further symbols. Suppose that **F** properly begins the formula **A** and that **E** properly ends the formula **A**. Show by formula induction that nl(**A**) = nr(**A**), that nl(**F**) = nr(**F**) = 0 or nl(**F**) > nr(**F**), and that nr(**E**) = nl(**E**) = 0 or nr(**E**) > nl(**E**).

b. Given an inductive definition of an occurrence of **B** in a formula **A** being a <u>syntactic</u> occurrence.

c. Using a., show that every occurrence of a formula within a formula is a syntactic occurrence.

21. (Unique Readability.) a. Say that the expression **E** <u>properly begins</u> a formula **A** if **A** consists of E followed by further symbols. Prove that no formula properly begins a formula **A**. (Hint: Use course-of-values formula induction on **A**. See section on induction in the appendix. It is important to understand each sentence letters as a single symbol or p_1 would begin p_{11} and the result would be untrue.)

b. Prove theorem 1.3. (Hint: Disjunctions begin with **(**, negations begin with ¬ and a sentence letter is comprised of one symbol, so disjunctions, negations and sentence letters comprise distinct classes. Use part a to establish the that the disjuncts of a disjunction are unique.)

c. Show that unique readability remains when left and right-hand brackets are replaced by the symbol |.

d. Show that the unique readability property remains when all of the right-hand brackets (or all of the left hand brackets) are omitted. (This observation is due to Ed Keenan. Formulas are harder to read under this convention, but the proof of unique readability is actually simpler than under the usual convention.)