CAMBRIDGE

Cambridge University Press
978-0-521-36770-7 - Systems of Logic
Norman M. Martin
Excerpt
More information

1
Formal Systems and Structure Theory

An essential characteristic of the systems which we are going to introduce is the fact that they are finite sequences of sets satisfying a certain principle of composition. The reader will note in the following that, usually, the particular nature of the sets which form the sequences is not essential. Specifically, we will define a **zero-order system** as an infinite family of decidable sets S_i ($i = 0,1,2,\cdots$). We will term the union of the S_i (i.e., the set of those elements that are elements of an S_i), S and we can formally define:

$$S = \bigcup S_i \ (i \in \mathbb{N})$$

To qualify as a zero-order system, two further conditions must be satisfied:

c1. If $i \neq j$, $S_i \cap S_j = \varnothing$ (i.e., nothing is in more than one S_i)

c2. If $x, y \in S$, $x \neq y$ and y is a sequence, then x is not an initial segment of y.

Depending on the details of the theory of classes we assume, c2 may be dispensable. Its function is to ensure that every finite sequence of elements of S is uniquely decomposable. If A, AB, BC and C were all elements of S, ABC could presumably be decomposed as AB followed by C or as A followed by BC. In some ways of interpreting classes, these sequences could be interpreted so that they are distinct; in such a case, c2 would presumably be redundant. However, in order to make our definition independent of the fine detail of our set theory, we adopt c2 nonetheless.

Let S' be the sequence $S_0, S_1, S_2, \cdots$. (More formally, S' is a set of ordered pairs whose first elements are sets and whose second are natural numbers such that if $< A,i > \in S'$ and $< B,i > \in S'$, then $A = B$ and distinct first elements satisfy c1 and c2.) We can now define $\sigma_{S\frown}$, the set of finite sequences of elements of S. We will say that a set T has property $\mathcal{T}_{S\frown}$ if and only if (iff) it satisfies the following conditions:

1. The null sequence, $\varnothing$, is an element of T ($\varnothing \in T$)
2. If there exists an i such that $x \in S_i$, then $x \in T$.
3. If $y \in T$ and $x \in S_i$, for some i, then there is a $z \in T$ such that $z =$

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-36770-7 - Systems of Logic
Norman M. Martin
Excerpt
More information

$x^\frown y$ (which we will normally write xy).

Then we can define $\sigma_{S\frown}$ (the set of finite sequences of elements of S or **formulae**) as:

$\sigma_{S\frown} = \cap\, T$ (T has the property $\mathcal{T}_{S\frown}$).

Strictly speaking, $\sigma_{S\frown}$ is a function of S and of the meaning assigned to "follows" ($\frown$); but only if $\frown$ satisfies the following conditions for all values:

F1. $\varnothing^\frown A = A^\frown\varnothing = A$

F2. If $A \neq B$, $A \neq \varnothing$, $B \neq \varnothing$, then $A^\frown B \neq B^\frown A$, for all A and B in S.

F3. $A^\frown(B^\frown C) = (A^\frown B)^\frown C$

F4. $A^\frown B^\frown C = A^\frown D^\frown C$, iff $B = D$

F5. If $A \subset S$, $B \subset S$ and $A \neq B$, then $A^\frown D \neq B^\frown C$.

The reader will note that many relations between many kinds of elements will satisfy F1-F5, but in our later text we will not specify which of these relations we use. Also that when, later on, we have occasion to speak of "sequence of sequences", "sequence of sequences of sequences", it is neither affirmed nor denied that the notion of "following" which is then used is in each case the same, but only that in each case F1-F5 are satisfied. We shall, in our examples, use linear (left to right) order of certain marks for first level sequences, a vertical ordering down the page for second order sequences, and additional columns at sufficient distance for third level ones, but these specific choices are only for the sake of convenience.

We now define a property C applicable to elements of $\sigma_{S\frown}$.

1. If $x \in S_0$, then x has property C.
2. If $f \in S_k$ $(k > 0)$ and $x_1,\cdots,x_k$ have property C, then $fx_1\cdots x_k$ has property C.

We then define W_S to be the intersection of all sets of elements of $\sigma_{S\frown}$ which have property C.

Following the terminology applied in the most familiar type of zero-order system, we will in the future call elements of S **symbols** and elements of W_S **well-formed formulae** (**wffs** for short). The reader should be warned that we have not used any of the properties of symbols in the ordinary sense (except that they can be members of certain sets), nor shall we, so that "symbol" and "wff" are being used in a very

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-36770-7 - Systems of Logic
Norman M. Martin
Excerpt
More information

abstract, perhaps peculiarly denatured sense.

In this connection, it is worth remarking that much of the customary terminology of modern logic is strongly affected by the linguistic views of some of the most prominent contributors to the field, such as Frege, Russell, and Carnap. In this book, we will not try either to defend these views or to subject them to thorough criticism. Nonetheless, the formulations of this book neither presuppose these views nor require that they be rejected. This happy situation (which we believe to be of some philosophical significance) is due largely to the somewhat abstract standpoint adopted. It does, however, have some unpleasant terminological consequences. Since the terminology commonly used is influenced by these linguistic views, we frequently have to choose between inventing new terms with the risk of being incomprehensible, or else using the customary terms with the risk of, at least, being misleading, and, at worst, reinforcing the Fregean-Logical Positivist family of views (by, so to speak, "brainwashing"). Since we regard these views as debatable rather than established, we would consider this to be unfortunate.

Faced with this choice, we have, as the reader will see, temporized. Where an alternative usage is available which is free of these philosophical implications, we have chosen it. Therefore, we will generally refer to "structure theory" and "proof theory" rather than "syntax", and also "model theory" rather than "semantics" (following Kleene). Similarly, we have chosen to use the term "first-order" instead of "lower predicate" or its variants with regard to calculi and logics and even, by analogy, "zero-order" instead of "propositional" or "sentential," to refer to those calculi whose only variables have wffs as substituends. (This characterization is used for identification; our formal definitions will come later.) On the other hand, we have preserved the common terms "symbol", "formula", "predicate," "variable," and the like, despite the fact that our definitions do not imply, as the terms themselves suggest, that they are linguistic or that they are symbolic in the Peirce-Mead sense of referring to something else. This philosophic semantic reticence should not be taken as insisting that one **cannot** so interpret them, but only that one need not do so. The reader should be careful when understanding our concepts -- and using them -- not to smuggle into them features suggested by the more usual meanings of the terms employed. We shall endeavor to give explicit warnings when this seems appropriate.

In the succeeding text, we shall be interested in characterizing in very general terms the elements or building blocks which go into the construction of a logical system (**structure theory**), the notions of

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-36770-7 - Systems of Logic
Norman M. Martin
Excerpt
More information

derivation and proof and the comparative relations between systems (**proof theory**), the notions of interpretation and satisfaction and the resulting notions of implication (or entailment) and equivalence (**model theory**), and the interrelations of the three. In addition, we will apply the resulting concepts in our examination of a number of logical systems, guided in our choice by several kinds of pragmatic motivations, which should be clear as we go along. Considerations of space and simplicity prevent our exploring more than a few possibilities. In our treatment, we shall be more thorough with regard to zero-order systems (and among them, the non-modal ones), although we shall include a partial treatment of some other systems as well. Unless specifically stated otherwise, omissions are dictated by space, convenience, and to a degree, our judgment as to what is easier to learn first and do not suggest that systems and topics not covered are unimportant.

Somewhat similar considerations of convenience have influenced us in our decision to omit first and higher order concepts from our initial presentation of structure and proof theory and to limit our first treatment of model theory to extensional methods. When in our later discussions the earlier presentations are used, we will say so.

For our presentation we will need several concepts from elementary mathematics. Suppose there is a property P of natural numbers such that:

α. 0 has property $\mathcal{P}$

β. For any number n, if n has property $\mathcal{P}$, then so does n+1; under these conditions, a form of reasoning which is termed **mathematical induction** (or, more specifically, **weak induction**) allows us to conclude

Every number has property $\mathcal{P}$.

We shall not attempt to prove this form of reasoning, but simply adopt it without further ado. For those who are not familiar with it and possibly worried by its adoption, it may be helpful to point out that each particular case subsumed under the general conclusion asserts that $\mathcal{P}$ holds for some particular finite number k. Premise α asserts $\mathcal{P}$ of 0. Then by β, $\mathcal{P}$ holds of 1. By the same argument, also of 2, and of 3, and of 4, and so on. But no matter how large k may be, if it is (as we have stated) a finite number, we can eventually count up to it, so that the argument will eventually get us to assert that k has property $\mathcal{P}$ (without specific use of induction).

An alternative form, frequently called **strong induction**, and equivalent to the preceding for finite numbers, consists of the premise:

For every finite number k, if n has property $\mathcal{P}$ for all numbers

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-36770-7 - Systems of Logic
Norman M. Martin
Excerpt
More information

n ($0 \leq n < k$), then k has property $\mathcal{P}$;

and the conclusion:

Every number has property $\mathcal{P}$.

We will, where convenient, use this form as well. We can of course present the same kind of argument as before. Drawing then a figurative deep breath, we return to the consideration of zero-order systems:

We now define a function on $\sigma_{S^\frown}$, whose values are natural numbers and which may be informally considered the **length of** or the **number of symbols in** an element of $\sigma_{S^\frown}$. The function $\ell(x)$ with the domain $\sigma_{S^\frown}$ is defined, as follows:

1. $\ell(\varnothing) = 0$
2. $\ell(xy) = 1 + \ell(y)$, provided $x \in S$

It follows trivially that $\ell(x) = 0$ if and only if $x = \varnothing$ and as a result, $\ell(x) = 1$ if and only if $x \in S$. Hence if neither x nor y is the empty sequence $\varnothing$, $\ell(xy) > \ell(x)$ and $\ell(xy) > \ell(y)$. Given sequences x and y, we say that x is an **initial segment** of y provided there exists a sequence z in $\sigma_{S^\frown}$ such that $y = xz$.

We will now prove a number of basic theorems:

Theorem 1-1. If x and y are in $\sigma_{S^\frown}$ then xy is in $\sigma_{S^\frown}$.

Proof: By induction on $\ell(x)$:

(α) $\ell(x) = 0$. Then $x = \varnothing$ by the definition of ℓ. Hence $xy = \varnothing y = y$, by F1. Since $xy = y$ and $y \in \sigma_{S^\frown}$, $xy \in \sigma_{S^\frown}$.

(β) Assume the theorem is true for all z with $z \in \sigma_{S^\frown}$ and $\ell(z) < k$. Suppose $\ell(x) = k > 0$. Then there exist v and w such that $v \in S$, $w \in \sigma_{S^\frown}$ and $x = vw$. Then $xy = (vw)y = v(wy)$, by F3 . Then $\ell(wy) = k\text{-}1$ and hence $wy \in \sigma_{S^\frown}$ and therefore $xy \in \sigma_{S^\frown}$.

Theorem 1-2. If $x_1,\cdots,x_k \in \sigma_{S^\frown}$, then so is $x_1\cdots x_k$.

Proof: By induction on k:

(α) For $k = 1$, there is nothing to prove

(β) Assume the theorem is true for $k = n$. We will prove it for

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-36770-7 - Systems of Logic
Norman M. Martin
Excerpt
More information

$k = n+1$. By the hypothesis of induction, $x_1 \cdots x_n \in \sigma_{S\cap}$. Hence, by theorem 1-1, $x_1 \cdots x_n x_{n+1} \in \sigma_{S\cap}$.

Theorem 1-3. Let x, y, z, w be elements of $\sigma_{S\cap}$ and $xz = yw$. Then either x is an initial segment of y or y of x.

Proof: By induction on $\ell(x)$:

(α) If $\ell(x) = 0$, then $x = \varnothing$ and hence $y = \varnothing y = xy$.

(β) Assume the theorem holds for all v with $\ell(v) < k$. Prove that it holds for x with $\ell(x) = k$. Since x and y are elements of $\sigma_{S\cap}$, there must exist $t \in S$ and $u \in \sigma_{S\cap}$ such that $x = tu$, and also $q \in S$ and $r \in \sigma_{S\cap}$ such that $y = qr$. Thus $(tu)z = (qr)w$. Hence by F3, $t(uz) = q(rw)$. By F5, it follows that $t = q$ and hence by F4, $uz = rw$. Since $\ell(u) < \ell(x)$, the induction hypothesis implies u is an initial segment of r, or r of u. Consequently tu is an initial segment of tr or vice versa and since $x = tu$ and $y = qr = tr$, x is an initial segment of y, or vice versa.

Theorem 1-4. If S is a zero-order system, $x \in W_S$ and $xy \in W_S$, then $y = \varnothing$.

Proof: By induction on $\ell(x)$:

(α) If $\ell(xy) > \ell(x)$ and $x \in W_S$, $\ell(x) \geq 1$. Let $\ell(x) = 1$. Then $x \in S_0$. If $xy \in S_0$, $xy = fz_1 \cdots z_k$ and $f \in S_k$. But then by c2, $x = f$, so that $k = 0$ and $y = \varnothing$. If, however, $xy \in S_0$, $\ell(xy) = 1$. Since $\ell(xy) = 1 + \ell(y)$, $\ell(y) = 0$. Thus $y = \varnothing$.

(β) Suppose the theorem is true for $\ell(x) < k^*$. We prove it for x' and y' such that $\ell(x') = k^* > 1$. Then $\ell(x'y') \geq \ell(x') > 1$. Hence $x' = fz_1 \cdots z_k$ with $f \in S_k$, $z_1, \cdots, z_k \in W_S$ and $x'y' = gy_1 \cdots y_j$ with $g \in S_j$ and $y_1, \cdots, y_j \in W_S$, then $x'y' = gy_1 \cdots y_j = fz_1 \cdots z_k y'$ and by c2, $f = g$ and hence $j = k$. Let i be the first index (if any) such that $z_i \neq y_i$. Then $z_i \cdots z_k y' = y_i \cdots y_k$. Since $\ell(z_i) < \ell(x') = k^*$ and by the hypothesis of induction z_i cannot be an initial segment of y_i. Since, however, z_i is an initial segment of $y_i \cdots y_k$ then y_i is an initial segment of z_i by theorem 1-3, but then $\ell(y_i) = \ell(z_i) < k^*$ and as

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-36770-7 - Systems of Logic
Norman M. Martin
Excerpt
More information

above y_i cannot be an initial segment of z_i. Hence $x' = fz_1 \cdots z_k = gy_1 \cdots y_k = x'y'$. Since $x'\varnothing\varnothing = x'y'\varnothing$, we have $y' = \varnothing$ by F4.

These results lead to the following rather far-reaching theorem:

Theorem 1-5. (**Unique Decomposition Theorem for Well-Formed Formulae**) Let $x_1, \cdots, x_j, y_1, \cdots, y_k \in W_S$, $f \in S_j$, $g \in S_k$, and $fx_1 \cdots x_j = gy_1 \cdots y_k$. Then $j = k$, $f = g$ and $x_i = y_i$ for $i = 1, \cdots, j$.

Proof: By induction on $\ell(fx_1 \cdots x_j)$:

(α) $\ell(fx_1 \cdots x_j) = 1$. Then $fx_1 \cdots x_j \in S_0$, $gy_1 \cdots y_k \in S_0$ and hence $j = k = 0$ and $f = g$.

(β) Assume the theorem is true for all x' such that $x' \in W_S$ and $\ell(x') < k^*$. Let $\ell(fx_1 \cdots x_j) = k^*$. By F5, $f = g$ and hence $j = k$ by C1. Let i be the least natural number such that $x_i \neq y_i$. If i exists, let $x_{i+1} \cdots x_k$ be X and $y_{i+1} \cdots y_k$ be Y. Then $x_iX = y_iY$ by F4 and hence by theorem 1-3, either x_i is an initial segment of y_i or vice versa. By theorem 1-4, neither can be the case. If i doesn't exist, there is nothing more to prove.

The Unique Decomposition Theorem guarantees the uniqueness of the way any wff can be decomposed into partial wffs. One consequence of it that will be useful is:

Theorem 1-6. Let $x, v \in W_S$, $y, z, w \in \sigma_{S\cap}$ and $xy = zvw$. Then either zv is an initial segment of x or x is an initial segment of z.

Proof: By theorem 1-3, either zv is an initial segment of x or x of zv. Let us assume that the former fails and hence the latter holds. Then we show that x is an initial segment of z. By induction on $\ell(x)$:

(α) Assume $\ell(x) = 1$. Suppose $z \neq \varnothing$. Hence there is a t and a u such that $t \in S$ and $u \in \sigma_{S\cap}$ and $z = tu$. Thus $xy = tuvw$ and $x = t$, and x is an initial segment of z. If $z = \varnothing$, $xy = vw$ and since by assumption v (which is zv) is not an initial segment of x, x is an initial segment of v by theorem 1-3. Hence there is a t in $\sigma_{S\cap}$ such that $v = xt$ and $t = \varnothing$ by theorem 1-4, so that $zv\varnothing = zv = v = x$ and

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-36770-7 - Systems of Logic
Norman M. Martin
Excerpt
More information

zv is, contrary to assumption, an initial segment of x.

(β) Assume the theorem holds for all $x' \in W_S$ such that $\ell(x') < k^*$. Assume $\ell(x) = k^* > 1$. Then there exist k, f and $x_1,\cdots,x_k$ such that $f \in Sk$ and $x_1,\cdots,x_k \in W_S$ and $x = fx_1\cdots x_k$ by theorem 1-5. Since x is an initial segment of zw there exists a t in $\sigma_{S\cap}$ such that $fx_1\cdots x_k t$ = zv. Hence by theorem 1-3, $fx_1\cdots x_k$ is an initial segment of z (in which case the proof is done) or z is an initial segment of $fx_1\cdots x_k$. Assume the latter. If $z = \varnothing$, x is an initial segment of v and hence by theorem 1-4, x = v = zv and therefore zv is an initial segment of x contrary to assumption. If $z \neq \varnothing$, there exist u and p such that $u \in S$, $p \in \sigma_{S\cap}$ and z = up. Then $fx_1\cdots x_k t = upv$ and hence f = u and $x_1\cdots x_k t = pv$. Let j be the least natural number such that $x_1\cdots x_j$ is not an initial segment of p. If j doesn't exist, then x is an initial segment of fp = up = z and the proof is done. Assume that j exists. Then $x_1\cdots x_{j-1}$ is an initial segment of p (when j = 0, $x_1\cdots x_{j-1}$ is taken to be $\varnothing$). Therefore, there is a $q \in \sigma_{S\cap}$ such that $x_1\cdots x_{j-1}q = p$. Hence, $x_1\cdots x_{j-1}\, x_j\cdots x_k t = x_1\cdots x_{j-1}qv$ and by F4, $x_j\cdots x_k t = qv$. It follows by theorem 1-3 that q is an initial segment of x_j. Since, however, $\ell(xj) < k^*$, and $qv = qv\varnothing$ and $xj \in W_S$ and $v \in W_S$, by the hypothesis of induction, qv is an initial segment of x, contrary to assumption, or xj is an initial segment of q contrary to assumption. It follows that no such j could exist and hence x is an initial segment of z.

We will now define an operator on W_S which is intended to produce the result of uniformly substituting one wff for another in a wff. It will be denoted $S^A{}_B C$ (to be read **the result of substituting B for A uniformly in C**) and is defined as follows:

1. $S^A{}_B A = B$
2. If $f \in S_k$, $k > 0$ and $x_1,\cdots,x_k \in W_S$, and $A \neq fx_1\cdots x_k$, then $S^A{}_B fx_1\cdots x_k = fS^A{}_B x_1\cdots S^A{}_B x_k$.
3. If $A \neq C$ and either $C \in S_0$ or $C \notin W_S$, then $S^A{}_B C = C$.

Theorem 1-7. If A, B and $C \in W_S$, $S^A{}_B C \in W_S$.

Proof: By induction on $\ell(C)$:

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-36770-7 - Systems of Logic
Norman M. Martin
Excerpt
More information

(α) If $\ell(C) = 1$, then $C \in S_0$. Hence either $A = C$ and $\mathbf{S}^A{}_B C = B$, or $A \neq C$ and $\mathbf{S}^A{}_B C = C$. Hence $\mathbf{S}^A{}_B C \in W_S$.

(β) Suppose the theorem true for all D such that $\ell(D) < k$ and assume that $\ell(C) = k > 1$. Since $C \in W_S$, there exist n, f and $x_1 \cdots x_n$ such that $C = fx_1 \cdots x_n$, $f \in S_n$ and $x_1 \cdots x_n \in W_S$. Hence $\mathbf{S}^A{}_B C = f\mathbf{S}^A{}_B x_1 \cdots \mathbf{S}^A{}_B x_n$. By the definition of ℓ, $\ell(x_i) < k$, for each i $(1 \leq i \leq n)$. By the hypothesis of induction, $\mathbf{S}^A{}_B x_i \in W_S$, and so is $\mathbf{S}^A{}_B C$.

To extend the notion of substitution to several wffs, we first define a **sequential substitution**:

$$\mathbb{S}^{A_1}{}_{B_1} C = \mathbf{S}^{A_1}{}_{B_1} C$$

$$\mathbb{S}^{A_1}{}_{B_1} \cdots {}^{A_{n+1}}{}_{B_{n+1}} C = \mathbf{S}^{A_{n+1}}{}_{B_{n+1}} \mathbb{S}^{A_1}{}_{B_1} \cdots {}^{A_n}{}_{B_n} C$$

We can then define **simultaneous substitution** as:

$$\mathbf{S}^{A_1}{}_{B_1} \cdots {}^{A_n}{}_{B_n} C = \mathbb{S}^{v_1}{}_{B_1} \cdots {}^{v_n}{}_{B_n} \mathbb{S}^{A_1}{}_{v_1} \cdots {}^{A_n}{}_{v_n} C$$

where $v_1, \cdots, v_n$ are distinct elements of S_0 that do not occur in any of $A_1, \cdots, A_n, B_1, \cdots, B_n$ or C.

Note that no insurmountable difficulty results even if a formal system $\mathfrak{S}$ has only a finite number of elements in S_0 and hence that the conditions on $v_1, \cdots, v_n$ may turn out to be unsatisfiable. This is true because we can, in that case, introduce a system $\mathfrak{S}'$ which differs from $\mathfrak{S}$ only by having an infinite set of additional elements of S_0, $\{w_1, w_2, \cdots\}$ that are not in the alphabet of $\mathfrak{S}$, and then define our simultaneous and sequential substitution operators in $\mathfrak{S}'$. Since no w_i will occur in $\mathbf{S}^{A_1}{}_{B_1} \cdots {}^{A_n}{}_{B_n} C$, the resulting operator always generates an element of W_S. (The reader who now worries about the existence of non-elements of S would have to be reminded that this language is recursive so that standard diagonalization arguments, unfortunately beyond the scope of this work, will guarantee their existence.) We can use the $\mathbf{S}$-operator to define a notion of occurrence for wffs as follows: **A S-occurs in B** iff A and B are wffs andeither:

1. If W_S contains at least two elements, there exists a $C \in W_S$

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-36770-7 - Systems of Logic
Norman M. Martin
Excerpt
More information

such that $\mathcal{S}^{A}{}_{C}B \neq B$, or

2. W_S has only one element

Theorem 1-8. If $B = fx_1 \cdots x_k \in W_S$, then A S-occurs in B, if and only if, either $A = B$ or for some i ($1 \leq i \leq k$), A S-occurs in x_i.

Proof: We assume A S-occurs in B. If $A = B$, there is nothing to prove, so we assume $A \neq B$. By the definition of S-occurs there is a $C \in W_S$ such that $\mathcal{S}^{A}{}_{C}B \neq B$. Since $\mathcal{S}^{A}{}_{C}B = f\mathcal{S}^{A}{}_{C}x_1 \cdots \mathcal{S}^{A}{}_{C}x_n$, there must be an i such that $\mathcal{S}^{A}{}_{C}x_i \neq x_i$. Hence A S-occurs in x_i. We now prove the implication in the other direction. For our first case, we assume $A = B$. If A is the only element of W_S, then A S-occurs in B by definition. Otherwise, there is a $C \in W_S$ such that $C \neq A$. Therefore, $\mathcal{S}^{A}{}_{C}B = \mathcal{S}^{A}{}_{C}A = C \neq A$ and we again have A S-occurs in B. Next we assume there exists an i ($1 \leq i \leq k$) such that A S-occurs in x_i. Then there exists a C with $\mathcal{S}^{A}{}_{C}x_i \neq x_i$. By theorem 1-5, $B \neq f\mathcal{S}^{A}{}_{C}x_1 \cdots \mathcal{S}^{A}{}_{C}x_n = \mathcal{S}^{A}{}_{C}B$. This implies that A S-occurs in B.

Theorem 1-9. If A S-occurs in $\mathcal{S}^{A}{}_{B}C$, $A \in S_0$, $B \in W_S$ and $C \in W_S$, then A S-occurs in B.

Proof: By induction on $\ell(C)$:

(α) Assume $\ell(C) = 1$. Then $C \in S_0$. If $C = A$, $\mathcal{S}^{A}{}_{B}C = B$ and hence A S-occurs in $\mathcal{S}^{A}{}_{B}C$ iff A S-occurs in B. If $C \neq A$, then $\mathcal{S}^{A}{}_{B}C = C$. Hence we also have that for every $D \in W_S$, $\mathcal{S}^{A}{}_{B}C = C$ and therefore $\mathcal{S}^{A}{}_{D}\mathcal{S}^{A}{}_{B}C = \mathcal{S}^{A}{}_{D}C = C$, contrary to the assumption that A S-occurs in $\mathcal{S}^{A}{}_{B}C$.

(β) Assume the theorem holds provided $\ell(C) < j$. We will prove it for $\ell(C) = j$. Since $C \in W_S$, there exist k, f and $x_1, \cdots, x_k$ with $f \in S_k$ and with the $x_i \in W_S$ such that $C = fx_1 \cdots x_k$. Then $\mathcal{S}^{A}{}_{B}C = f\mathcal{S}^{A}{}_{B}x_1 \cdots \mathcal{S}^{A}{}_{B}x_n$. If A S-occurs in $\mathcal{S}^{A}{}_{B}C$, either $A = \mathcal{S}^{A}{}_{B}C$ or there is an i ($1 \leq i \leq k$) such that A S-occurs in $\mathcal{S}^{A}{}_{B}x_i$ by theorem 1-8. If the former, $A = f$ and hence $f \in S_0$ so that $\ell(C) = 1$, contrary to assumption. Thus A must

© Cambridge University Press www.cambridge.org