Notre Dame Journal of Formal Logic
Volume XII, Number 4, October 1971
NDJFAM

A GENERALIZATION OF THE GENTZEN HAUPTSATZ

LUIS E. SANCHIS

1. Gentzen rules describe proofs in first order logic on the basis of simple transformations, each one related to the meaning of some logical operation, which are enough to avoid the use of the cut rule. A fundamental application of the rules is the reduction of predicate logic to propositional logic via the so-called extended Gentzen Hauptsatz or midsequent theorem (also called the Herbrand-Gentzen theorem in [1]). We present in this paper another reduction which seems very convenient for consistency proofs of universal theories. It is shown that if such a theory is inconsistent then some inconsistency can be found just using the cut rule to eliminate atomic formulas.

In the system considered by Gentzen only one kind of axiom is admitted: those of the form $A \to A$. We allow other axioms and prove the elimination of the cut rule under certain restrictions. The resulting system provides a good frame for the description of axiomatic theories and preserves the characteristic symmetry of the Gentzen rules.

2. We shall deal with some given first order language in which we assume the following symbols: free individual variables, bound individual variables, individual constants, function letters, the equality symbol, predicate letters, propositional connectives: $-$, $\vee$, $\wedge$, $\supset$, $\equiv$, and quantifiers: $\forall$, $\exists$.

Letters $a, b, c, \ldots$ are used as syntactic variables for free individual variables and letters $x, y, z, \ldots$ for bound individual variables. Terms are defined in the usual way. Atomic formulas are either expressions of the form $R(t_1, \ldots, t_n)$ where R is some predicate letter and $t_1, \ldots, t_n$ are terms, or expressions of the form $t = h$ where t and h are terms. Formulas are defined by induction in the usual way.

Letters $M, N, P, \ldots$ will denote finite (possible empty) sequences of formulas. The formulas of the sequence M are called components of M. We introduce a symbol $\to$ and expressions of the form $M \to N$ are called sequents.

The notation $A(t/b)$ denotes the result obtained when the term t is substituted for the variable b in the formula A and it is also a formula. If

Received February 23, 1970

M is the sequence $A_1, \ldots, A_n$ then $M(t/b)$ denotes the sequence $A_1(t/b), \ldots, A_n(t/b)$.

If all the components of a sequence M are atomic we say that M is atomic. If both M and N are atomic we say that the sequent $M \to N$ is atomic.

The notation $M < N$ means that every component of M is also a component of N and in that case we say that N is an expansion of M. Hence if M and N are considered as sets (disregarding order and repetitions) this notation indicates ordinary set inclusion. If P is an expansion of M and Q is an expansion of N then we say that $P \to Q$ is an expansion of $M \to N$.

We list below the derivation rules. These rules describe transformation from one or two sequents (called the premises of the rule and written above a line) to another sequent called the conclusion of the rule and written below the line. In these rules M, N, P, Q are arbitrary sequences, A and B are arbitrary formulas, t is any term, b is any free variable, c is a free variable that does not occur in M or N, and x is a bound variable that does not occur in A.

$$\text{Expansion: } \frac{M \to N}{P \to Q} \text{ provided } P \to Q \text{ is an expansion of } M \to N$$

$$\text{- left: } \frac{M \to N, A}{-A, M \to N} \qquad \text{- right: } \frac{A, M \to N}{M \to N, -A}$$

$$\wedge \text{ left: } \frac{A, B, M \to N}{(A \wedge B)\ M \to N} \qquad \wedge \text{ right: } \frac{M \to N, A \quad P \to Q, B}{M, P \to N, Q, (A \wedge B)}$$

$$\vee \text{ left: } \frac{A, M \to N \quad B, P \to Q}{(A \vee B), M, P \to N, Q} \qquad \vee \text{ right: } \frac{M \to N, A, B}{M \to N, (A \vee B)}$$

$$\supset \text{ left: } \frac{M \to N, A \quad B, P \to Q}{(A \supset B), M, P \to N, Q} \qquad \supset \text{ right: } \frac{A, M \to N, B}{M \to N, (A \supset B)}$$

$$\equiv \text{ left: } \frac{A, B, M \to N \quad P \to Q, A, B}{(A \equiv B), M, P \to N, Q} \qquad \equiv \text{ right: } \frac{A, M \to N, B \quad B, M \to N, A}{M, P \to N, Q\ (A \equiv B)}$$

$$\forall \text{ left: } \frac{A(t/b), M \to N}{\forall x\, A(x/b), M \to N} \qquad \forall \text{ right: } \frac{M \to N, A}{M \to N, \forall x A\,(x/c)}$$

$$\exists \text{ left: } \frac{A, M \to N}{\exists x A(x/c), M \to N} \qquad \exists \text{ right: } \frac{M \to N, A(t/b)}{M \to N, \exists x A\,(x/b)}$$

$$\text{Cut: } \frac{M \to N, A \quad A, P \to Q}{M, P \to N, Q}$$

A basic set E is a set of atomic sequents that contains all sequents $A \to A$ for any atomic formula A. We say that E is closed if, whenever $M \to N$ is in E then $M(t/b) \to N(t/b)$ is also in E for any term t and free variable b. Let E be some basic set. We shall say that a sequent $M \to N$ is E-valid if it can be obtained by a finite number of applications of the derivations rules starting with sequents in E. Hence $M \to N$ is E-valid if and only if it belongs to the smallest set that contains E and is closed under the derivations rules. We shall say that $M \rightarrow N$ is strictly E-valid if it can be obtained from E without using the cut rule. If $M \rightarrow N$ is E-valid there is a derivation from E which can be described in a tree form as explained in [2, p. 106]. We call such tree an E-derivation tree. The notion of the height of a tree is used as in [2]. If the cut rule is not used in the derivation tree we say it is a strict E-derivation tree.

Lemma 1. *Let E be closed and suppose there is a (strict) E-derivation tree for $M \rightarrow N$ of height k. Then, given a term t and a variable a there is a (strict) E-derivation for $M(t/a) \rightarrow N(t/a)$ of height k.*

The proof is by induction on k. In case of rules $\forall$ right and $\exists$ left the induction hypothesis allows changing the variables in the proof in such a way that the restrictions for those rules are satisfied.

3. We consider in this section the possibility of eliminating the cut rules in E-derivations. In general this is not possible but it is easy to show that the cut rule can be restricted to the elimination of atomic formulas.

If M is a sequence and A is a formula then $M(A)$ denotes the sequence obtained by deleting in M all components identical with A. The following rule can be obtained by means of the cut and expansion rules (where A is any formula that appears as a component of both N and P):

$$\text{Mix rule: } \frac{M \rightarrow N \quad P \rightarrow Q}{M,\, P(A) \rightarrow N(A), Q}$$

Given a basic set E, we call E^* the smallest set that contains E and is closed under the mix rule. Clearly E^* is also a basic set. If E is closed then E^* is also closed.

Theorem 1. *Let E be a closed basic set and suppose $M \rightarrow N$ and $P \rightarrow Q$ are both strictly E^*-valid. Assume $N < S, A$ and $P < T, A$ for sequences S and T and formula A. Then $M, T \rightarrow S, Q$ is also strictly E^*-valid.*

This theorem is essentially the elimination of the cut rule for E^*-derivations. Gentzen's proof of the Hauptsatz actually reduces the complex cases of the cut rule to elimination between axioms. In our system the closure property of E^* will take care of that situation. Hence it is clear that Gentzen's proof can be adapted to our theorem. We have chosen a formulation of the theorem which is very convenient in dealing with the different cases of the induction. We do not think it is necessary to consider these cases in detail since the ideas are the same as appear in the proofs of the Hauptsatz in [1] or [2], but we shall explain the kind of induction which is used in the proof.

Let k be the number of connectives and quantifiers in the formula A. Suppose there is a strict E^*-derivation tree for $M \rightarrow N$ of height m and another strict E^*-derivation tree for $P \rightarrow Q$ of height n. Then we use a main induction on k and a secondary induction on $m + n$. The cases $m = 1$ and $n = 1$ are covered by the closure property of E^*. In the other cases it can be assumed that either $M \rightarrow N$ or $P \rightarrow Q$ is obtained in the tree by the application of some rule. Whenever the formula introduced in one of these rules is not A the secondary induction hypothesis will be sufficient. If the formula A is introduced in both rules then it is necessary to use the main induction hypothesis.

Corollary. *Let E be a closed basic set. Then, if $M \to N$ is E^*-valid, it is also strictly E^*-valid.*

Theorem 2. *Let E be a closed basic set. Then, the following conditions are equivalent:*
(i) $M \to N$ is E-valid
(ii) $M \to N$ is strictly E^*-valid
(iii) there is some E-derivation for $M \to N$ in which for all applications of the cut rule the formula eliminated is atomic.

If $M \to N$ is E-valid then it is also E^*-valid, hence by the corollary to Theorem 1 it is strictly E^*-valid.

If $M \to N$ is strictly E^*-valid then there is an E^*-derivation in which the cut rule is not used. But any sequent which is not in E but is in E^* can be obtained from E by means of the mix rule, hence using the cut rule and expansions. In this way we obtain an E-derivation in which for all applications of the cut rule, the formula eliminated is atomic.

That (*iii*) implies (*i*) is trivial.

Remark. Let E_0 be the set of all atomic sequents of the form $A \to A$ where A is an atomic formula. Then E_0 is closed and $E_0^* = E_0$. In this case Theorem 2 reduces to the ordinary Gentzen Hauptsatz.

Theorem 3. *Let E be a closed basic set and $M \to N$ an atomic sequent. Then $M \to N$ is E-valid if and only if it is the expansion of some sequent in E^*.*

Suppose $M \to N$ is E-valid. Then by Theorem 2 there is a strict E^*-derivation for $M \to N$. In this derivation the only rule that can be used is the expansion rule, hence $M \to N$ is the expansion of some sequent in E^*. The converse is trivial.

Corollary. *The empty sequent is E-valid if and only if E^* contains the empty sequent.*

For these results the condition that E is closed is essential. For instance, consider the set E that contains exactly two sequents: $\to A$, and $A(t/b) \to$, where A is some atomic formula, b is a variable that occurs in A and t is a term different from b. The empty sequent is E-valid, but $E^* = E$, and hence the empty sequent is not in E^*.

4. We shall say that a basic set is consistent if the empty sequent is not E-valid. In this section we give a semantical characterization of such sets E for the case when E is closed. We know that in this case E is consistent if and only if E^* does not contain the empty sequent.

We shall consider assignments of truth values to the atomic formulas of the language. A given assignment satisfies a sequent $M \to N$ if either there is a component of N that takes the value true, or there is a component of M that takes the value false. Hence, the empty sequent is never satisfied. An assignment satisfies a basic set E if it satisfies all the sequents in E. Hence, if E is empty, then it is satisfied by all assignments.

We say that an atomic sequent $M \to N$ is a consequence of the basic set E if, whenever some assignment satisfies E then it satisfies $M \to N$ also.

A tautology is an atomic sequent which is satisfied by all assignments. It is clear that $M \to N$ is a tautology if and only if M and N have some common component.

Theorem 4. *Let E be some basic set and $M \to N$ an atomic sequent. The following conditions are then equivalent:*

(*i*) *$M \to N$ is a consequence of E.*

(*ii*) *$M \to N$ is an expansion of some sequent in E^*.*

It is clear that (*ii*) implies (*i*). To prove the converse note that if $M \to N$ is a tautology then it is an expansion of some sequent in E, since E is a basic set. So we assume that $M \to N$ is not a tautology and is not an expansion of some sequent in E^* and prove that $M \to N$ is not a consequence of E.

Let $A_1, A_2, \ldots$ be an enumeration without repetitions of all atomic formulas in the language excluding those that are components of M or N. We define a complete string as an infinite sequence of sequents: $M_0 \to N_0$, $M_1 \to N_1, \ldots$ such that $M_0 \to N_0$ is $M \to N$ and $M_{k+1} \to N_{k+1}$ is either $A_k, M_k \to N_k$ or $M_k \to N_k, A_k$, for $k = 0, 1, \ldots$ If both $A_k, M_k \to N_k$ and $M_k \to N_k, A_k$ are expansions of some sequent in E^* then $M_k \to N_k$ is also an expansion of some sequent in E^*. Hence there is one complete string with the property that no sequent in the string is the expansion of some sequent in E^*. Let L be one such complete string.

It is clear that every atomic formula appears as a component in some sequent of L. Since $M \to N$ is not a tautology there is no atomic formula that appears as a component both in the left and the right side of some sequent of L. Hence, we use L to define the following assignment of truth values: an atomic formula which is a component in the left side of some sequent in L takes the value true and any atomic formula appearing as component in the right takes the value false. This assignment satisfies all sequents in E^*, for if some sequent is not satisfied this means that after some step all sequents in the complete string are expansions of such a sequent. Since $M \to N$ is not satisfied by the assignment it follows that it is not a consequence of E.

Corollary. *E^* does not contain the empty sequent if and only if there is some assignment that satisfies E.*

Theorem 5. *Let E be closed. Then, E is consistent if and only if it is satisfiable by some assignment.*

Remark. If T is a theory with equality in which all the axioms are universal sentences, then there is a closed basic set E such that for any formula A, A is a theorem of T if and only if A is E-valid. For instance, suppose a language is given with just one unary function letter f and a binary predicate letter R and the only axiom is the formula:

$$\forall x \forall y\ (R(x, f(y)) \equiv (R(x, y) \vee x = y))$$

Then, E consists of all sequents of the following form for arbitrary terms t, h, g, t_1, t_2, h_1, h_2:

$$R(t, f(h)) \rightarrow R(t, h),\ t = h$$
$$R(t, h) \rightarrow R(t, f(h))$$
$$t = h \rightarrow R(t, f(h))$$
$$\rightarrow t = t$$
$$t = h \rightarrow h = t$$
$$t = h,\ h = g \rightarrow t = g$$
$$t = h \rightarrow f(t) = f(h)$$
$$R(t_1, h_1),\ t_1 = t_2,\ h_1 = h_2 \rightarrow R(t_2, h_2)$$

For theories in which not all axioms are universal sentences it is necessary to eliminate the existential quantifiers by the use of Skolem functions to get a similar basic set E.

REFERENCES

[1] Curry, H. B., *Foundations of Mathematical Logic,* McGraw-Hill, New York (1963).

[2] Kleene, S. C., *Introduction to Metamathematics,* Van Nostrand, Princeton, New Jersey (1950).

Pennsylavania State University
University Park, Pennsylvania