In: Proceedings of the Workshop on Programming Logic, Båstad, Sweden, 21-26 May 1989. Ed. by Peter Dybjer, Lars Hallnäs, Bengt Nordström, Kent Petersson and Jan M. Smith. Programming Methodology Group, University of Göteborg and Chalmers University of Technology 1989, Report 54, pp. 494-517.

Judgements of Higher Levels and Standardized Rules for Logical Constants in Martin-Löf's Theory of Logic

Peter Schroeder-Heister

The aim of these notes is to carry over some of what I did in my thesis to the framework of Martin-Löf's logical theory, in particular the idea of rules of higher levels (which in Martin-Löf's non-formalistic approach will become hypothetical judgements of higher levels) and the general schema for introduction and elimination rules for logical constants (which will have to be extended to a schema containing formation and detraction rules). I make no claims to originality. Concerning Martin-Löf's system I mainly rely on his Siena lectures of 1983. In the first part, I shall deal with propositional logic, and in the second part I shall try to show how the results extend to logical constants of any arity, leaving out, however, the theory of expressions which is an integral part of Martin-Löf's logical theory, but which is not immediately necessary for the understanding of the logical rules.

I. Propositional Logic

a) Categorical and hypothetical judgements

Propositional logic is that part of logic which deals with closed expressions and certain n-ary constants to be defined as logical operators. According to Martin-Löf, it does not deal with "propositions" which are given as a domain of discourse from outside. Whether a closed expression is a proposition is something that is to be established within the theory. Otherwise the theory would loose its formal character ("formal" = independent of the content)

or would become a formal system ("formal" = syntactically defined) which is interpreted from outside.
Therefore Martin-Löf's theory distinguishes two forms of categorical judgements, A is a proposition (A prop) and A is true (A true) which are explained in such a way that the latter presupposes the former.

More precisely, A prop and A true are explained by telling what it means to <u>know</u> them, i.e. what it means to have proved them. Proof, as well as judgement, is understood as an act, not as a formal object. A proof of a judgement is the act which makes this judgement evident, i.e. known, to someone. So judgements are not justified independently of the subject who makes the judgement. This does not mean that formal proofs are no longer allowed. Once I have seen that a certain inference step leads me from a judgement which is evident to me to another one which is evident to me, I can later on use this step as a formal rule of inference, relying on the evidence for this step which I had and which I can reproduce if I want. However, the basic concept of proof with respect to which formal rules are justified, is the subject-dependent one.

The explanations of A prop and A true run as follows:

To know A prop means to know what one must do in order to verify A, i.e. what counts as a verification of A. So if I have grasped what a verification of A looks like, I have proved A prop. For example, if I know the procedure which would verify an observation statement A, I have proved A prop. It is obvious that this diverges from the usual notion of proof. The <u>explanation</u> of a verification procedure, provided it is understood, already constitutes the <u>proof</u> of a judgement. (There is no dichotomy in principle between explanation and demonstration.)

A true is explained only under the presupposition that A prop has been proved. So suppose A prop is known. Then to know A true means to know how to verify A, i.e. to be able to produce a verification of A, i.e. to produce something of which one knows what it looks like because of the presupposition A prop.

According to these explanations, A prop is always a judgement (perhaps an unjustified one), whereas A true is a judgement only under the condition that A prop has been proved. This has to do with what Martin-Löf calls the intentional character of propositions, following Heyting and Kolmogorov. If propositions are intentions, and verifications are fulfilments of intentions, then in order to be in a position to verify A one must first know what counts as a verification of A, since an intention (even if it is not successful) is possible only on the basis of knowledge of what is intended.

Note that to verify A is not the same as to prove A true. Otherwise A true could not have been explained by reference to verification. Verification is a basic notion which is used to express the intentionality connected with propositions. Of course, if I have verified A, I know how to verify A and thus have proved A true. But conversely, if I have proved A true, I only know <u>how</u> to verify A; therefore the further step of executing this knowledge is necessary to obtain a verification.

Similar to A true, most cases of hypothetical judgements will be explained under the presupposition that certain other judgements (which are already explained) have been proved. In the following, if R is to be explained as a judgement, by $\mathcal{D}$(R) I shall denote those judgements which must have been explained before and which are supposed to be known (= proved). So in general an explanation of R will take the form:

(*) { Suppose $\mathfrak{D}(R)$ has been proved. Then to know (= to have proved) R means ....

In the case A prop, $\mathfrak{D}$(A prop) is empty (so there is no presupposition), and the dots are to be replaced by the above explanation of A prop. In the case A true, $\mathfrak{D}$(A true) is A prop, and the dots are to be replaced by the above explanation of A true.

Furthermore we introduce the following terminology: We call a candidate R for the explanation as a judgement a <u>potential judgement</u>. A potential judgement R is called a judgement if $\mathfrak{D}(R)$ has been proved. This terminology is justified by the fact that if $\mathfrak{D}(R)$ has been proved, then the explanation (*) of R can be applied to R saying what it means to know R, i.e. explaining R as a judgement. For example, A prop is a potential judgement for any A and at the same time a judgement since $\mathfrak{D}$(A prop) is empty, i.e. A prop is explained without any precondition. A true is a potential judgement for any A and a judgement if $\mathfrak{D}$(A true) (= A prop) has been proved, for then the explanation of A true can be applied.

The full apparatus of hypothetical judgements is introduced by the following definitions: A prop and A true are potential judgements for any closed expression A. If $R_1,\dots,R_n$,R are potential judgements, then so is $(R_1,\dots,R_n)\Rightarrow R$. Potential judgements different from A prop and A true are potential hypothetical judgements. Lists of potential judgements are of the form $(R_1,\dots,R_n)$ or $\emptyset$ (empty list), the R_i being called elements if the list. U, V, W, X, Y, Z denote lists of potential judgements, R and R' (with and without indices) potential judgements. (X,Y) or (X,R) etc. are understood as usual. In the notation of lists we usually omit outer brackts. Single potential judgements are considered to be limiting cases of lists. Our task now is to define $\mathcal{D}(R)$ and to give an explanation of the form (*) for each R.

The main idea is that to know $(R_1,\dots,R_n)\Rightarrow R$ means to have a hypothetical proof of R from $R_1,\dots,R_n$, which in turn means that one has a proof of R which is uniform in $R_1,\dots,R_n$. "Uniform in $R_1,\dots,R_n$" means that after supplementation by proofs of the $R_1,\dots,R_n$ one immediately obtains a proof of R (proof here understood in the primary, categorical sense). To have a hypothetical proof of R from $R_1,\dots,R_n$ constitutes a new single act of knowledge, it is not an infinite collection of proofs of R, one for each list of proofs of the $R_1,\dots,R_n$. This is why I used the term "uniform". "Schematic" would be another term to express this fact. Hypothetical proofs result from categorical proofs by a similar kind of abstraction as do general proofs (i.e. proofs with free variables). Assumptions in hypothetical proofs can be viewed like variables to be instantiated by their proofs.

It is an essential feature of Martin-Löf's system that the supplementation of $R_1,\dots,R_n$ in a hypothetical proof of R from $R_1,\dots$ R_n by proofs of $R_1,\dots,R_n$ is not considered to be performed in one step, but may be done stepwise, i.e. by first supplementing R_1 by its proof, then supplementing R_2 by its proof, and so on. This makes a great difference in the presupposition under which $(R_1,\dots,R_n)\Rightarrow R$ is explained. According to the approach where the $R_1,\dots,R_n$ are considered to be replaced by proofs simultaneously, one would require (1) that all $R_1,\dots,R_n$ be judgements, and (2) that R be a judgement provided all $R_1,\dots,R_n$ have been proved (since it is not until the $R_1,\dots,R_n$ have been proved that R needs to be explained as a judgement). So $\mathcal{D}((R_1,\dots,R_n)\Rightarrow R)$ would be defined as $\mathcal{D}(R_1),\dots,\mathcal{D}(R_n),(R_1,\dots,R_n)\Rightarrow\mathcal{D}(R)$. According to Martin-Löf's approach, R_2 need not be explained as a judgement until R_1 has been proved, R_3 not until R_1 and R_2 have been proved, etc. This leads to the following definition of $\mathcal{D}$, where $\mathcal{D}$ assigns lists of potential judgements to lists of potential judgements:

$$\mathcal{D}(A \text{ prop}) = \emptyset$$

$$\mathcal{D}(A \text{ true}) = A \text{ prop}$$

$$\mathcal{D}(X \Rightarrow R) = \mathcal{D}(X), X \Rightarrow \mathcal{D}(R) \quad (\text{not } (\mathcal{D}(X), X) \Rightarrow \mathcal{D}(R) \text{ !})$$

$$\mathcal{D}(X, R) = \mathcal{D}(X), X \Rightarrow \mathcal{D}(R).$$

Here the implicit convention is used that $X \Rightarrow \emptyset$ is identified with $\emptyset$ and $X \Rightarrow (R_1, \ldots, R_n)$ with the list $X \Rightarrow R_1, \ldots, X \Rightarrow R_n$. For example,

$\mathcal{D}((A_1 \text{ true}, A_2 \text{ true}, A_3 \text{ true}) \Rightarrow A \text{ prop}) = \mathcal{D}(A_1 \text{ true}, A_2 \text{ true}, A_3 \text{ true})$

$= \mathcal{D}(A_1 \text{ true}, A_2 \text{ true}), (A_1 \text{ true}, A_2 \text{ true}) \Rightarrow A_3 \text{ prop}$

$= \mathcal{D}(A_1 \text{ true}), A_1 \text{ true} \Rightarrow A_2 \text{ prop}, (A_1 \text{ true}, A_2 \text{ true}) \Rightarrow A_3 \text{ prop}$

$= A_1 \text{ prop}, A_1 \text{ true} \Rightarrow A_2 \text{ prop}, (A_1 \text{ true}, A_2 \text{ true}) \Rightarrow A_3 \text{ prop}.$

For $\mathcal{D}((A_1 \text{ true}, A_2 \text{ true}, A_3 \text{ true}) \Rightarrow A \text{ true})$ one would have to add $(A_1 \text{ true}, A_2 \text{ true}, A_3 \text{ true}) \Rightarrow A \text{ prop}$.

It follows from the definition of $\mathcal{D}$ that $\mathcal{D}((X, Y) \Rightarrow R) = \mathcal{D}(X \Rightarrow (Y \Rightarrow R))$, which is quite natural and which would not hold in the non-stepwise conception, save one would define $\mathcal{D}(R)$ to be just the list of those A prop for which A true occurs in R. This would make a potential hypothetical judgement a hypothetical judgement only if it is built up from categorical judgements.

$\mathcal{D}$ has the following property:

Lemma 1 If R' is an element of $\mathcal{D}(R)$, then each element of $\mathcal{D}(R')$ is an element of $\mathcal{D}(R)$. That is, elementwise application of $\mathcal{D}$ to $\mathcal{D}(R)$ does not yield anything new (whereas $\mathcal{D}(\mathcal{D}(X))$, which is not defined elementwise, can yield something new).
[I do not reproduce any proofs of lemmas or theorems in these notes.]

Now the precise explanation of a hypothetical judgement $(R_1, \ldots, R_n) \Rightarrow R$ is the following: Suppose all elements of $\mathcal{D}((R_1, \ldots, R_n) \Rightarrow R)$ have been proved. Then to know $(R_1, \ldots, R_n) \Rightarrow R$

means to have a hypothetical proof of R from $R_1,\ldots,R_n$ in the sense that it becomes a proof of R by stepwise supplementation by proofs of $R_1,\ldots,R_n$.

We must convince ourselves that this is a genuine explanation. If we measure the complexity of a potential judgement R' by the pair (number of occurrences of A true in R', number of occurences of A prop in R'), then each element of $\mathcal{D}(R')$ is of lower complexity than R'. Furthermore, according to Lemma 1, $\mathcal{D}(R')$ contains all presuppositions of the explanations of its elements, so that it makes sense to assume all of them to be proved. Thus, when ordered according to their complexity, all potential hypothetical judgements are covered by the explanation in a non-circular way. Moreover, because of the presupposition of the explanation, it can be considered explained what it means to have proofs of $R_1,\ldots,R_n$ and of R depending on $R_1,\ldots,R_n$: Since $\mathcal{D}(R_1)$ belongs to $\mathcal{D}((R_1,\ldots,R_n)\Rightarrow R)$, R_1 is explained. Since for each $i<n$, $(R_1,\ldots,R_i)\Rightarrow\mathcal{D}(R_{i+1})$ belongs to $\mathcal{D}((R_1,\ldots,R_n)\Rightarrow R)$, R_{i+1} is explained provided $R_1,\ldots,R_i$ have been proved. That is, the $R_1,\ldots,R_n$ can be considered explained step by step, provided in each step the previous judgements have been proved, which is all that must be required for a stepwise supplementation by proofs of $R_1,\ldots,R_n$. Since $(R_1,\ldots,R_n)\Rightarrow\mathcal{D}(R)$ belongs to $\mathcal{D}((R_1,\ldots,R_n)\Rightarrow R$ R is explained provided $R_1,\ldots,R_n$ have been proved which is all that is necessary for a hypothetical proof of R from $R_1,\ldots,R_n$.

Therefore the above explanation, together with the explanations of A prop and A true, explains each R as a judgement provided $\mathcal{D}(R)$ has been proved. Hence we can, as already proposed, call R a judgement if $\mathcal{D}(R)$ has been proved. As an extension of this mode of speach, we shall call a list of potential judgements X a <u>system of judgements</u>, if $\mathcal{D}(X)$ (i.e. each element of $\mathcal{D}(X)$) has been proved. In a system of judgements, each element is a judgement provided the previous elements have been proved. In Martin-Löf's framework, the assumptions of hypothetical proofs are systems of judgements in this sense (for which order is important), and not just finite sets.

Theorem 1 For any list of potential judgements X, $\mathcal{D}(X)$ is a system of judgements, and therefore also $\mathcal{D}(X),X$.

b) General rules of inference

Rules of inference lead one from judgements which are known (= proved to another judgement which is known. So the justification of a rule of inference must show that if I have proofs of the premisses I also obtain a proof of the conclusion. Now the proofs in question may themselves be hypothetical proofs, and it is useful to make the assumptions of such hypothetical proofs explicit using the notation X:R for a hypothetical proof of R from X (where Ø:R denotes the limiting case of a categorical proof). So the general form of a rule of inference is

$$\frac{X_1:R_1 \quad \dots \quad X_n:R_n}{X:R},$$

to be read as: if hypothetical proofs of $R_1,\dots,R_n$ from $X_1,\dots,X_n$, respectively, are given, one obtains a hypothetical proof of R from X. Unlike Martin-Löf's, this notation also mentions those assumptions which are not discharged by the application of the rule. For example, instead of

$$\frac{\begin{matrix}(A \text{ true})\\ B \text{ true}\end{matrix}}{A \supset B \text{ true}}$$

we write

$$\frac{X,\ A \text{ true}: B \text{ true}}{X : A \supset B \text{ true}}\ .$$

The reason is that in the present framework the assumptions of a hypothetical proof form a more complicated structure than in usual natural deduction proofs. In particular, order is important, as the example

$$\frac{A \text{ true}, X : B \text{ true}}{X : A \supset B \text{ true}}$$

shows, which is not in general a valid rule of inference.

The proposed notation allows one to formulate rules of inference whose justification is based only on the explanations of categorical and hypothetical judgements and is still independent of the introduction of logical constants. I call these rules, which correspond to Gentzen's structural rules, general rules of inference, in contradistinction to special rules of inference which govern the logical constants.

Note that I use the colon as a specific sign to express hypothetical proofs which is different from $\Rightarrow$, i.e., I distinguish a categorical proof of $X \Rightarrow R$ (expressed by $\emptyset : X \Rightarrow R$) from a hypothetical proof of R from X (expressed by X:R). This is justified since the notion of a hypothetical proof is the primary notion with respect to which the notion of a hypothetical judgement is explained. Establishing $X \Rightarrow R$, given a hypothetical proof of R from X, is an extra inference step, even if it is an immediate one (based on the explanation of hypothetical judgements).

However, this is not a matter of principle. Everything that follows remains valid if one replaces the colon by $\Rightarrow$. The difference is that when using the colon a rule of inference is conceived as something that leads one from hypothetical proofs to a hypothetical proof (where something may be changed in the assumptions, e.g. assumptions may be discharged), whereas when using $\Rightarrow$, a rule is viewed as leading from hypothetical judgements (or categorical proofs thereof) to hypothetical judgements (or categorical proofs thereof). Which interpretation one prefers depends on whether one wants to make the step of reflection, which lies between proofs of hypothetical judgements and hypothetical proofs, explicit in the notation of rules of inference.

In the following formulation of the general rules of inference we use the convention that X:Y means $X{:}R_1 \ \dots \ X{:}R_n$ if Y is the list $R_1,\dots,R_n$, and that X:Y is empty if Y is empty.

$$(\mathrm{Contr})\ \frac{X,R,Y,R \ :\ R'}{X,R,Y \ :\ R'} \qquad\qquad (\mathrm{Thin})\ \frac{X,Y \ :\ R' \qquad X \ :\ \mathcal{D}(R)}{X,R,Y \ :\ R'}$$

$$(\mathrm{Ass})\ \frac{X \ :\ \mathcal{D}(R)}{X,R \ :\ R} \qquad \text{(where X must be empty if } \mathcal{D}(R) \text{ is empty)}$$

$$(\mathrm{Hyp})\ \frac{X \ :\ (R_1,\dots,R_n)\Rightarrow R \qquad X \ :\ R_1 \ \ \dots \ \ X \ :\ R_n}{X \ :\ R}$$

$$(\Rightarrow)\ \frac{X,Y \ :\ R}{X \ :\ Y\Rightarrow R}$$

Note that according to the formulation of (Ass), R can be introduces as an assumption only if R is a judgement (i.e., $\mathcal{D}(R)$ has been proved). Hence R:R is not in general justified. Since A prop is always a judgement (i.e. $\mathcal{D}(\text{A prop}) = \emptyset$), A prop can always be assumed, i.e. A prop : A prop is justified.

These rules can be justified in the following sense.

<u>Theorem 2</u> Let

$$\frac{X_1{:}R_1 \ \dots \ X_n{:}R_n}{X{:}R}$$

be one of the rules in question. If for each premiss, (X_i,R_i) is a system of judgements (i.e., $\mathcal{D}(X_i,R_i)$ has been proved) and a hypothetical proof of R_i from X_i is given, then one obtains a proof of $\mathfrak{D}(X,R)$ (i.e., (X,R) is a system of judgements) and a hypothetical proof of R from X.

This theorem is proved by reflection on the explanation of the forms of judgement. [Again, I omit the detailed proof here]. Roughly speaking, the theorem says that if the premisses are explained and proved then so is the conclusion. The justification of a rule of inference includes showing that the conclusion is explained before showing that it is proved, since to speak of a hypothetical proof of R from X makes sense only if (X,R) is a system of judgements, i.e. if a proof of $\mathfrak{D}(X,R)$ is given. The latter may depend itself on the proofs of the premisses of the rule.

Having justified certain rules of inference, we may consider the formal calculus we obtain if we take the general rules of inference to be formal rules which allow one to produce sequences of signs from sequences of signs already produced. In that case, we shall speak of formal provability and formal proofs, as distinguished from proofs as acts which make something evident. The formal calculus has proof-theoretic properties which correspond in a certain sense to properties which have to do with non-formal proofs. This is not surprizing since the justification of the general rules of inference may be viewed as a demonstration of the soundness of the corresponding formal system.

Theorem 3 Consider the calculus based on (Contr), (Thin), (Ass), (Hyp) and ($\Rightarrow$) as formal rules of inference.

(i) For each X, both $\mathfrak{D}(X) : \mathfrak{D}(X)$ and $\mathfrak{D}(X), X : X$ are formally provable.

(ii) If X:R is formally provable, then $\emptyset : \mathfrak{D}(X,R)$ is formally provable.

(i) corresponds to Theorem 1, (ii) in part to Theorem 2. Note that from (i) we may conclude that there are proofs in the non-formal sense of the elements of $\mathfrak{D}(X)$ from $\mathfrak{D}(X)$ and from the elements of X from $\mathfrak{D}(X)$,X. Just take the formal proofs and combine the justifications of the inference rules which are used. (ii) does not have an immediate non-formal reading, since for having a hypothetical proof of R from X (in the non-formal sense) it is already presupposed that a proof of $\mathfrak{D}(X,R)$ is at one's disposal. The significance of (ii) lies in the fact that, though it is possible to speak of a formal proof of X:R without any presupposition, the presupposition of the non-formal case has a formal analogue in the formal provability of $\emptyset$: $\mathfrak{D}(X,R)$. If we had not made assumptions explicit and thus had not formulated the general rules of inference, we would at least have lost this nice correspondence between results of non-formal reflections and proofs in a certain formal system. The structural rules of this formal system are not as simple as in the case of ordinary natural deduction, where explicit mentioning of assumptions and structural rules can well be avoided as in Gentzen's first and in Prawitz's presentation.

There are of couse further general rules of inference which can be justified, e.g. the following ones:

$$(\text{Ass}')\ \frac{X : \mathfrak{D}(Y)}{X,Y : Y} \quad \text{(Where X must be empty if } \mathfrak{D}(Y) \text{ is empty)}$$

$$(\text{Thin}')\ \frac{X,Y : R \qquad X : \mathfrak{D}(Z)}{X,Z,Y : R} \qquad (\Rightarrow')\ \frac{X : Y \Rightarrow R}{X,Y : R}$$

$$(\text{Contr}')\ \frac{X,R,Y,R,Z : R'}{X,R,Y,Z : R'} \qquad (\text{Perm})\ \frac{X,R_1,R_2,Y : R \qquad X : \mathfrak{D}(R_2)}{X,R_2,R_1,Y : R} .$$

These rules, read as formal rules, can be shown to be admissible in the calculus based on (Contr), (Thin), (Ass), (Hyp), ($\Rightarrow$), i.e. do not extend what is formally provable in this calculus.

c) Logical operators - special rules of inference

Now we consider n-ary constants S and explain their meaning by telling what counts as a verification of $SA_1 \ldots A_n$. More precisely, we do not give an explanation for specific constants but give a schema of an explanation for an arbitrary n-ary constant S, which, when instantiated appropriately, becomes an explanation of the specific logical constants one wants to have, e.g. &, v, $\supset$ and $\perp$. In general this explanation will depend on the assumption that something has been proved which must have been explained before. This will be expressed by use of lists of potential judgements to be associated with S.

Let $p_1, \ldots, p_n$ be additional closed expressions, called propositional variables, and let $\Delta_1(p_1, \ldots, p_n), \ldots, \Delta_m(p_1, \ldots, p_n)$ be lists of potential judgements associated with S, whose expressions are built up only by use of propositional variables and logical constants which have already been explained (if there are any), and which contain potential categorical judgements of the form "A true" only. The lists may be empty, and we even allow for m to be 0, in which case no list (not even the empty one) is associated with S. $\Delta_i(A_1, \ldots, A_n)$ is obtained from $\Delta_i(p_1, \ldots, p_n)$ by simultaneously substituting $A_1, \ldots, A_n$ for $p_1, \ldots, p_n$, respectively. I shall also write $\overline{p}$ for $p_1, \ldots, p_n$, $\overline{A}$ for $A_1, \ldots, A_n$, $S\overline{A}$ for $SA_1 \ldots A_n$, $\Delta_i(\overline{p})$ for $\Delta_i(p_1, \ldots, p_n)$, and $\Delta_i(\overline{A})$ for $\Delta_i(A_1, \ldots, A_n)$.

Now the meaning of S is explained as follows: Let $\overline{A}$ be given. Suppose $\Delta_i(\overline{A})$ is a system of judgements for every i. Then a verification of $S\overline{A}$ consists of a proof of (the elements of) $\Delta_i(\overline{A})$ for some i.

Let us use $\{—_j—\}_j$ for $—_1— \dots —_m—$ (empty if m=0), where $—_j—$ is anything containing an index j, and $\{—_j—\}_{i \neq j}$ for $—_1— \dots —_{i-1}— —_{i+1}— \dots —_m—$. Then the explanation of S immediately leads to the following rule of inference:

$$(\text{S form})\ \frac{\{X : \mathcal{D}(\Delta_j(\bar{A}))\}_j}{X : S\bar{A}\ \text{prop}} \quad .$$

<u>Justification</u>: If one knows that $(X, \mathcal{D}(\Delta_j(\bar{A})))$ is a system of judgements for every j, one knows in particular that X is a system of judgements and thus that $(X, S\bar{A}$ prop) is a system of judgements. So the conclusion is explained. If one furthermore has a hypothetical proof of $\mathcal{D}(\Delta_j(\bar{A}))$ from X for every j, then the presupposition of the explanation of $S\bar{A}$ is fulfilled (provided X). Thus it is explained what counts as a verification of $S\bar{A}$, i.e. one knows $S\bar{A}$ prop (provided X). This is exactly what the conclusion of the rule says.

The inverses of the formation rule are the detraction rules (the idea of introducing detraction rules was developed jointly with Roy Dyckhoff, and the term "detraction rules" is due to him).

$$(\text{S detr})\ \frac{X : S\bar{A}\ \text{prop}}{\{X : \mathcal{D}(\Delta_j(\bar{A}))\}_j}$$

(This is considered to be a list of rules in the obvious way.)

<u>Justification</u>: (Here and in the following I omit reference to X, which is a list of assumptions common to premiss and conclusion). If one knows $S\bar{A}$ prop, one knows what counts as a verification of $S\bar{A}$, thus one has grasped the explanation of $S\bar{A}$, which means that one must have proved $\mathcal{D}(\Delta_i(\bar{A}))$ for every i (which is the presupposition of the explanation). This is exactly what the conclusion asserts. (We need not show in addition that the conclusion is explained, i.e. that for each R in $\mathcal{D}(\Delta_i(\bar{A}))$ we have a hypothetical proof of $\mathcal{D}(R)$ from X. By Lemma 1, this is contained in what we have shown.)

Martin-Löf does not formulate detraction rules, probably because he does not need them in the development of his theory. My reason for the formulation of these rules is that without them certain rules of inference would not be equivalent in the formal reading although they are equivalent in the non-formal reading which shows that in non-formal reasoning detraction rules are used implicitly. This applies, for example, to the equivalence between direct and indirect elimination rules for operators with only one associated system $\Delta_1(\overline{p})$ (see below) and to the following two kinds of introduction rules.

$$\text{(S intr)}\ \frac{X : \Delta_i(\overline{A}) \qquad X : S\overline{A}\ \text{prop}}{X : S\overline{A}\ \text{true}} \qquad (1 \leq i \leq m)$$

$$\text{(S intr')}\ \frac{X : \Delta_i(\overline{A}) \qquad \{X : \mathcal{D}(\Delta_j(\overline{A}))\}_{j \neq i}}{X : S\overline{A}\ \text{true}} \qquad (1 \leq i \leq m)$$

In the presence of (S form) and (S detr) these rules are formally interderivable (for the proof one has to use Theorem 3(ii), which extends to the calculus with (S form) and (S detr)); without (S detr), (S intr) is the stronger rule. Since we have already justified (S form), it suffices to justify (S intr).

<u>Justification</u>: If I know $S\overline{A}$ prop, $S\overline{A}$ true is explained as a judgement. According to this explanation, a proof of $\Delta_i(\overline{A})$ is a verification of $S\overline{A}$. Since I have such a proof, I know how to verify $S\overline{A}$.

$$\text{(S elim)}\ \frac{X : S\overline{A}\ \text{true} \qquad \{X, \Delta_j(\overline{A}) : R\}_j}{X : R}$$

$$\text{(S elim')}\ \begin{cases} \dfrac{X : S\overline{A}\ \text{true} \quad \{X, \Delta_j(\overline{A}) : C\ \text{prop}\}_j}{X : C\ \text{prop}} \\[2ex] \dfrac{X : S\overline{A}\ \text{true} \quad \{X, \Delta_j(\overline{A}) : C\ \text{true}\}_j}{X : C\ \text{true}} \end{cases}$$

(S elim') is just an instance of (S elim). Conversely, (S elim) can be formally proved from (S elim') using $(\Rightarrow')$, (Perm), (Thin) and Theorem 3(ii). I justify (S elim).

Justification: If one knows $S\overline{A}$ true, one is able to produce a verification of $S\overline{A}$ which according to the explanation of $S\overline{A}$ consists of a proof of $\Delta_i(\overline{A})$ for some specific i. Together with the proof of R from $\Delta_i(\overline{A})$ (and X) and its presupposition, a proof of $\mathcal{D}(R)$ from $\Delta_i(\overline{A})$ (and X), one obtains proofs of $\mathcal{D}(R)$ and R (from X).

The elimination rules which Martin-Löf formulates follow the pattern

$$\frac{X : S\overline{A} \text{ true} \qquad X, S\overline{A} \text{ true} : C \text{ prop} \qquad \{X, \Delta_j(\overline{A}) : C \text{ true}\}_j}{X : C \text{ true}},$$

and seem to me to be too weak for some purposes. For example, the following theorem would not hold with them:

Theorem 4 If one adds to the formal calculus considered in Theorem 3 (S form), (S detr), (S intr) and (S elim) as formal rules of inference, then Theorem 3(ii) remains valid, i.e., if $X:R$ is formally provable, then so is $\emptyset : \mathcal{D}(X,R)$.

If only one list $\Delta(\overline{p})$ is associated with S, we formulate the following alternative introduction and elimination rules:

$$(S\ intr^*)\ \frac{X : \Delta(\overline{A})}{X : S\overline{A} \text{ true}}$$

$$(S\ elim^*)\ \frac{X : S\overline{A} \text{ true}}{X : \Delta(\overline{A})}.$$

The equivalence between (S intr*) and (S intr) is obvious. To prove that (S elim) and (S elim*) are formally interderivable, one must have detraction rules at one's disposal. Otherwise one cannot formally prove $X: \mathcal{D}(\Delta(\overline{A}))$, i.e. show that the conclusion is explained. For example,

$$\frac{X : A\&B \text{ true}}{X : A \text{ true}}$$

is formally shown to follow from (& elim) in the following way, where

$$\begin{array}{c} \vert \\ X : A\&B \text{ true} \end{array}$$

represents an arbitrary formal proof of X : A&B true which is supposed to be given:

$$\dfrac{\begin{array}{c} \vert \\ X : A\&B \text{ true} \end{array} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\begin{array}{c} \vert \\ X : A\&B \text{ true} \end{array}}{X : A\&B \text{ prop}}(\text{Theorem } 4)}{X : A \text{ prop}}(\&\text{ detr})}{X, A \text{ true} : A \text{ true}}(\text{Ass}) \qquad \dfrac{\dfrac{\begin{array}{c} \vert \\ X : A\&B \text{ true} \end{array}}{X : A\&B \text{ prop}}(\text{Theorem } 4)}{X, A \text{ true} : B \text{ prop}}(\&\text{ detr})}{X, A \text{ true}, B \text{ true} : A \text{ true}}(\text{Thin})}{X : A \text{ true}}(\&\text{ elim})$$

.

Functional completeness of &, v, $\supset$, $\bot$ is proved as in my thesis. This proof uses replacement of (B_1 true, ... , B_n true)$\Rightarrow$B tru by (B_1&...&B_n)$\supset$B true, so it crucially depends on the fact that the $\Delta_i(\bar{A})$ do not contain any B prop. The reason for not permitting B prop in $\Delta_i(\bar{A})$ is the intentional character of propositions. In the explanation of $S\bar{A}$, the fulfilment of the intention $S\bar{A}$ (= verification of $S\bar{A}$) is reduced to the fulfilment of the intentions $A_1,\ldots,A_n$, or more precisely, to certain relations between such fulfilments (expressed by hypothetical judgements of certain forms). And since only knowledge of B true leads, when executed, to the fulfilment of the intention B, we cannot permit B prop to occur in $\Delta_i(\bar{A})$. Knowledge of B prop does not lead to the fulfilment of an intention, but is only the presupposition which is necessary to understand B as an intention (whose realizability is established by a proof of B true). (In this sense, proofs of A prop only have an auxiliary function.)

This point is of extreme importance. If $\Delta(\overline{A})$ may contain B prop, one could define a one-place constant P with the associated list $\Delta(p)$ = p prop. This would lead to the rules

$$(\text{P form}) \frac{}{\emptyset : \text{PA prop}} \qquad (\text{P intr}) \frac{X : \text{A prop}}{X : \text{PA true}}$$

$$(\text{P elim}) \frac{X : \text{PA true}}{X : \text{A true}} \quad ,$$

which would represent an internal definition of the category of propositions. In an extended system this leads to a contradiction as shown by Aczel.

II. Quantifier Logic

a) General proofs and general judgements

Hypothetical judgements were explained using the notion of a hypothetical proof. A hypothetical proof of R from $R_1,\ldots,R_m$ has the characteristic feature that, when successively applied to proofs of $R_1,\ldots,R_m$, it becomes a proof of R. Similarly the notion of a general proof can be defined, if we consider expressions with free variables. If $R_1,\ldots,R_m$ and R contain no other free variables than $x_1,\ldots,x_n$, then a general proof of R from $R_1,\ldots,R_m$ (which in fact is a hypothetico-general proof, if the $R_1,\ldots,R_m$ are actually present,) is defined as something that becomes a proof of $R(^{x_1\ldots x_n}/_{A_1\ldots A_n})$ from $R_1(^{x_1\ldots x_n}/_{A_1\ldots A_n}),\ldots,$ $R_m(^{x_1\ldots x_n}/_{A_1\ldots A_n})$, if expressions A_i of the same arities as x_i are given. Here $(^{x_1\ldots x_n}/_{A_1\ldots A_n})$ means simultaneous substitution of the A_i for the corresponding x_i. This means that the proof is uniform or schematic in the x_i, i.e., we do not have an infinite collection of proofs (one for each list $A_1,\ldots,A_n$), but one single act of knowledge.

In the following, when speaking of a proof of R from X, this is to be understood in the sense of a general proof if X and R contain free variables. It can easily be seen that none of the general rules of inference justified in part I need be changed when read in this way. The explanation of general proofs and of free variables which serve to express the generality of proofs immediately justifies the following general rule of inference:

$$(\text{Subst})\ \frac{X : R}{X(^x/_A) : R(^x/_A)} \quad .$$

In order to introduce general judgements, we take as an additional clause in the definition of potential judgements the following: If R is a potential judgement, then so is $\Rightarrow_x R$ for a variable x. $\Rightarrow_{x_1} \ldots \Rightarrow_{x_n} (X \Rightarrow R)$ may be written as $X \Rightarrow_{x_1 \ldots x_n} R$. The definition of $\mathcal{D}$ is extended by $\mathcal{D}(\Rightarrow_x R) = \Rightarrow_x \mathcal{D}(R)$ (where, if $\mathcal{D}(R)$ has more than one element, this is taken elementwise). General judgements are then explained as follows: Suppose all elements of $\mathcal{D}(\Rightarrow_x R)$ have been proved. Then to know $\Rightarrow_x R$ means to have a general proof of R. This justifies the following general rules of inference:

$$(\text{Gen})\ \frac{X : R}{X : \Rightarrow_x R} \quad (x \text{ not free in } X)$$

$$(\text{Spec})\ \frac{X : \Rightarrow_x R}{X : R} \quad .$$

(Subst) and (Spec) could be formulated in one rule. However, taking different rules seems to me to be conceptually clearer. (Subst) has to do with the notion of a (hypothetico-) general proof, (Gen) and (Spec) with the notion of a (hypothetico-) general judgement as defined from this notion.

Similar to what was remarked in part I, we distinguish general proofs (X : R) from proofs of general judgements ($\emptyset : X \Rightarrow_{x_1 \ldots x_n} R$, where $x_1, \ldots, x_n$ are the free variables of X and R). If one does not want to draw this distinction, one may throughout replace X:R by $X \Rightarrow_{x_1 \ldots x_n} R$.

b) Logical constants

We immediately deal with higher-order logic, i.e. with logical constants of arbitrary arity. Here o is the arity of closed expressions and $(\alpha_1, \ldots, \alpha_n)$ the arity of expressions which, when applied to expressions of arities $\alpha_1, \ldots, \alpha_n$, yield a closed expression. The arity n of propositional logic now corresponds to $(\underbrace{o, \ldots, o}_{n \text{ times}})$. It is a characteristic feature of Martin-Löf's system that on the level of expressions it distinguishes only arities, starting from the one basic arity of closed expressions. Whether something belongs to a certain type or category cannot be learned from an inspection of the expression but is a judgement of the theory. This makes it very closely related to Frege's system, whereas in presentations of simple type theory, for example, it is usual to categorize expressions from the very beginning, in particular to start with two basic types of expressions, one for propositions or sentences, and one for individuals or terms.

So we assume that we have variables of any arity at our disposal, furthermore constants if we want. Simultaneous substitution $A(^{x_1 \ldots x_n}/_{B_1 \ldots B_n})$ of variables x_i by expressions B_i of corresponding arities in expressions A is defined as usual, similarly for potential judgements and lists of potential judgements. We assume that substitution includes relabelling of bound variables in such a way that it is always defined. Furthermore we assume that we have abstraction in the sense that $((x_1 \ldots x_n)A)$ is of arity $(\alpha_1, \ldots, \alpha_n)$, if A is of arity o and each x_i of arity α_i, and application in the sense that $CB_1 \ldots B_n$ is of arity o if C is of arity $(\alpha_1,\ldots,\alpha_n)$ and each B_i of arity α_i. If the arities of A, C, x_i and B_i are as above, $((x_1\ldots x_n)A)B_1\ldots B_n$ is considered definitionally equal to $A(^{x_1\ldots x_n}/_{B_1\ldots B_n})$ and $((x_1\ldots x_n)Cx_1\ldots x_n)$ to C. This justifies

$$(\beta\eta)\ \frac{X : R}{X : R'}$$

as a general rule of inference, where R' results from R by exchanging definitionally equal expressions in the above sense. In a thorough treatment of definitional equality between expressions, $(\beta\eta)$ would be reduced to more basic rules. However, for the present purposes, where we are mainly interested in hypothetical and general judgements and rules for logical constants based on them, it is enough to have $(\beta\eta)$.

The above explanation of (hypothetico-) general proofs and judgements and the justification of the inference rules based thereon held for any arity and was not confined to the arity of closed expressions.

Now we sketch how to deal with logical constants S which may be of arbitrary arity. The usual $\forall$ and $\exists$ quantifiers are considered to be of arity ((o)), $\forall xA$ and $\exists xA$ being abbreviations of $\forall((x)A)$ and $\exists((x)A)$. Propositional operators fit into the present framework as limiting cases.

If S is of arity $(\alpha_1,\ldots,\alpha_n)$, let $x_1,\ldots,x_n$ be distinguished variables of arities $\alpha_1,\ldots,\alpha_n$, respectively. As in part I, we shall use the abbreviation $\overline{x}$ for $x_1\ldots x_n$, and similarly $\overline{A}$ for $A_1\ldots A_n$ where the elements of $\overline{A}$ must correspond in arities to the elements of $\overline{x}$. Let again lists $\Delta_1(\overline{x}),\ldots,\Delta_m(\overline{x})$ of potential judgements be associated with S, which besides free variables of $\overline{x}$ may contain bound variables different from $\overline{x}$, additional free variables, and logical constants which have already been explained, but no other constants. As in the propositional case,'...prop' must not occur in them. For example, in the case of $\forall$ and $\exists$, m is 1 and $\Delta_1(x_1)$ is $\Rightarrow_y x_1 y$ and $x_1 y$, respectively, where x_1 is of arity (o) and y of arity o. $\Delta_i(\overline{A})$ is not defined as $(\Delta_i(\overline{x}))(\overline{x}/\overline{A})$, but as $(\Delta_i(\overline{x}))(\overline{x}\overline{y}/\overline{A}\overline{z})$, where $\overline{y}$ are the free variables in $\Delta_i(\overline{x})$ beyond $\overline{x}$, and $\overline{z}$ are distinct variables of the same arities which do not occur in $\overline{A}$. Thus $\Delta_i(\overline{A})$ includes relabelling of the additional free variables in $\Delta_i(\overline{x})$ in order to avoid confusion with the free variables in $\overline{A}$.

Now S is explained as follows: Let $\overline{A}$ be given. Let $\overline{z}$ consist of all variables which are free in at least one $\Delta_i(\overline{A})$ $(1 \leq i \leq m)$ but not free in $\overline{A}$. Suppose $\Rightarrow_{\overline{z}} \Delta_i(\overline{A})$ is a system of judgements for every i, that is, $\Rightarrow_{\overline{z}} \mathcal{D}(\Delta_i(\overline{A}))$ has been proved. Then a verification of $S\overline{A}$ consists of a proof of $(\Delta_i(\overline{A}))(\overline{z}/\overline{C})$ for some i and list of expressions $\overline{C}$.

It would not suffice just to require that $\Delta_i(\overline{A})$ be a system of judgements, i.e. to leave $\overline{z}$ as free variables. For if $S\overline{A}$ is to be proved hypothetically from X, this explanation itself is to be understood under the assumption X, and X may already contain some variable of $\overline{z}$ free.

I just state the special rules of inference - the justifications are completely along the lines of part Ic), only generality has to be taken into account in the obvious way.

$$(\text{S form})\ \frac{\{X : \Rightarrow_{\overline{z}} \mathcal{D}(\Delta_j(\overline{A}))\}_j}{X : S\overline{A} \text{ prop}}$$

$$(\text{S detr})\ \frac{X : S\overline{A} \text{ prop}}{\{X : \Rightarrow_{\overline{z}} \mathcal{D}(\Delta_j(\overline{A}))\}_j}$$

$$(\text{S intr})\ \frac{X : (\Delta_i(\overline{A}))(\overline{z}/\overline{C}) \qquad X : S\overline{A} \text{ prop}}{X : S\overline{A} \text{ true}} \quad (1 \leq i \leq m)$$

Note that if $\overline{z}$ is not empty (i.e. in a proper quantifier case), the following rule is not adequate:

$$(\text{S intr}')\ \frac{X : (\Delta_i(\overline{A}))(^{\overline{z}}/_{\overline{C}}) \qquad \{X : \Rightarrow_{\overline{z}} \mathcal{D}(\Delta_j(\overline{A}))\}_{j \neq i}}{X : S\overline{A}\ \text{true}} \quad (1 \leq i \leq m)$$

If the left premiss is explained, we have a hypothetical proof of $\mathcal{D}((\Delta_i(\overline{A}))(^{\overline{z}}/_{\overline{C}}))$ from X, but not necessarily of $\Rightarrow_{\overline{z}} \mathcal{D}(\Delta_i(\overline{A}))$ from X, which would be required to guarantee, together with the right premisses, a hypothetical proof of $S\overline{A}$ prop from X.

$$(\text{S elim})\ \frac{X : S\overline{A}\ \text{true} \qquad \{X, \Delta_j(\overline{A}) : R\}_j}{X : R} \quad (\overline{z}\ \text{not free in X or R})$$

$$(\text{S elim}')\begin{cases} \dfrac{X : S\overline{A}\ \text{true} \qquad \{X, \Delta_j(\overline{A}) : C\ \text{prop}\}_j}{X : C\ \text{prop}} \\[2ex] \dfrac{X : S\overline{A}\ \text{true} \qquad \{X, \Delta_j(\overline{A}) : C\ \text{true}\}_j}{X : C\ \text{true}} \end{cases} \quad (\overline{z}\ \text{not free in X or C})$$

If only one $\Delta(\overline{x})$ is associated with S, we have direct elimination rules only if $\Delta(\overline{x})$ contains no free variables beyond $\overline{x}$ (for example, $\exists$ has no direct elimination rules). In that case (S intr*) and (S elim*) are to be formulated as in the propositional case (the $\overline{A}$ now being expressions of possibly higher arity).

Functional completeness of &, v, $\supset$, $\perp$, $\forall$, $\exists$ can now be proved as in my Aachen paper. The quantifiers come in by translating general judgements in $\Delta_i(\overline{A})$ by the universal quantifier and free variables besides those in $\overline{A}$ by the existential quantifier.

Since we have immediately dealt with higher-order logic, one may ask whether Prawitz's result about the definability of logical operators in terms of $\forall$ and $\supset$ can be obtained in this framework. The answer is negative. For example, suppose we know AvB prop. Then we cannot prove

$$(\text{A true} \Rightarrow \text{x true, B true} \Rightarrow \text{x true}) \Rightarrow_x \text{x true}$$

from AvB true which would be needed for Prawitz's result. What one can prove from AvB true is

$$\text{x prop} \Rightarrow_x ((\text{A true} \Rightarrow \text{x true, B true} \Rightarrow \text{x true}) \Rightarrow \text{x true}),$$

or, if one uses $\Rightarrow_p R$ as an abbreviation for x prop $\Rightarrow_x R$,

$$(\text{A true} \Rightarrow \text{p true, B true} \Rightarrow \text{p true}) \Rightarrow_p \text{p true.}$$

This shows that Prawitz's result crucially depends on the fact that one considers propositionally restricted quantification to be a logical operation. In Martin-Löf's framework this is not permitted since '...prop' must not occur in the $\Delta_i(\overline{x})$ associated with logical operators. By use of propositionally restricted quantification one could internally define the category of propositions which would lead to a contradiction. What remains of Prawitz's result in Martin-Löf's framework is that if one knows $S\overline{A}$ prop, then one can prove

$$\{\Delta_j(\overline{A}) \Rightarrow \text{p true}\}_j \Rightarrow_p \text{p true}$$

from $S\overline{A}$ true and vice versa, but not necessarily $C(\overline{A})$ true from $S\overline{A}$ true and vice versa for some $C(\overline{x})$ for which $S\overline{x}$ prop $\Rightarrow_x C(\overline{x})$ prop holds. This shows once again that in Martin-Löf's system not every hypothetical judgement of higher level can be translated into a categorical judgement, so that hypothetical judgements of higher levels may have useful applications beyond questions of a standardized schema for elimination rules and functional completeness.

Peter Schroeder-Heister
June 1985