Ultimate Normal Forms for Parallelized Natural Deductions

NEIL TENNANT, *Department of Philosophy, The Ohio State University, 230 North Oval Mall, Columbus, OH 43210, USA.*
E-mail: tennant.9@osu.edu

Abstract

The system of natural deduction that originated with Gentzen (1934–5), and for which Prawitz (1965) proved a normalization theorem, is re-cast so that all elimination rules are in parallel form. This enables one to prove a very exigent normalization theorem. The normal forms that it provides have all disjunction-eliminations as low as possible, and have no major premisses for eliminations standing as conclusions of *any* rules. Normal natural deductions are *isomorphic* to cut-free, weakening-free sequent proofs. This form of normalization theorem renders unnecessary Gentzen's resort to sequent calculi in order to establish the desired metalogical properties of his logical system.

Ultimate normal forms are well-adapted to the needs of the computational logician, affording valuable constraints on proof-search. They also provide an analysis of deductive relevance. There is a deep isomorphism between natural deductions and sequent proofs in the relevantized system.

Keywords: cut-elimination, Hauptsatz, intuitionistic relevant logic, relevance, natural deduction, normalization, parallelized elimination rules, reduction procedure, sequent calculus, ultimate normal form

1 Introduction and historical background

1.1 Main aim

Our main aim is to prove the *Hauptsatz* directly for the intuitionistic system of natural deduction with *parallelized* elimination rules.

This is something Gentzen failed to do. His system of natural deduction did not have fully parallelized elimination rules. He resorted also to a detour through the sequent calculus. *Cut-free sequent proofs* became his formal explication of the informal idea of *proofs without detours*.

The classic source of normalization theorems for natural deduction, Prawitz [6], did not treat of fully parallelized elimination rules. Even more recent treatments of parallelized systems of natural deduction, such as that of Negri and von Plato [4], do not prove normalization *directly*. They still involve a detour through the sequent calculus. (We discuss this at greater length in §3.1 below.)

In our own earlier treatment in [17] (henceforth: *Autologic*), we began with normal-form Gentzen–Prawitz-style natural deductions in *minimal* logic, and showed how to convert them into what we here call 'ultimate normal form'. We did not begin with parallelized natural deductions. We did not treat the intuitionistic case. And we did not establish the deep isomorphism that obtains, in the relevantized fragment, between natural deductions in ultimate normal form and sequent proofs that use only logical rules. The present paper therefore represents an advance over our own earlier work.

We claim as an expository advantage that we deal, throughout, directly with actual proof-schemata, and avoid all use of λ-terms and typing. Our theoretical aims, focused as they are upon an analysis of relevance, are fully served by establishing a weak normalization theorem for the full system of parallelized intuitionistic logic with $\neg$, $\wedge$, $\vee$ and $\rightarrow$ primitive. We leave to others the proof, if they so desire, of a strong normalization theorem. We conjecture that strong normalization holds for the full parallelized system. (See [2] §6.6 for a proof of strong normalization of the parallelized intuitionistic system with $\rightarrow$ as its only logical operator.)

We shall offer here a different explication than Gentzen did of the informal idea of 'proofs without detours': *natural deductions in ultimate normal form*. The natural deductions in question use parallelized rules; and this new kind of normal form is very exigent. We shall prove the ultimate normalizability of natural deductions *directly*, remaining wholly within the setting of natural deduction. Our analysis reveals a direct isomorphism between natural deductions in ultimate normal form, and cut-free, weakening-free sequent proofs. *This* is why it does, directly, for the system of natural deduction exactly what Gentzen did with his detour into sequent calculus. There will be a valuable spin-off: an analysis of deductive *relevance* (of premisses to conclusion). This leads to the identification of the 'relevant fragments' of both classical and intuitionistic logic. These are enormously useful for the specification of algorithms for automated proof-search. The central result, to be established below, strengthens our result in [19], and is as follows:

Main Theorem. Any intuitionistic natural deduction that uses parallelized elimination rules and derives the conclusion A from the set X of undischarged assumptions can be transformed into a natural deduction, in intuitionistic relevant logic, whose conclusion is either A or $\perp$, whose undischarged assumptions are in X, and in which every elimination has its major premiss standing as an assumption.

1.2 Some notational and terminological preliminaries

We shall use the following abbreviations for various systems of logic: *C* for classical logic, *I* for intuitionistic logic, *M* for minimal logic, *CR* for classical relevant logic, and *IR* for intuitionistic relevant logic.

We shall restrict our investigation to systems of propositional logic, although the treatment generalizes to first-order logic. Thus we shall be concerned only with the logical connectives $\neg$ (negation), $\wedge$ (conjunction), $\vee$ (disjunction) and $\rightarrow$ (implication). The symbol $\perp$ (for absurdity) will not be a formula of the language; instead, it will occur only as a marker within proofs.

Definition: *Sequents* are of the form $X{:}Y$, where X and Y are (possibly empty) finite sets of sentences, and Y is at most a singleton. When we write $X, A{:}Y$, or $Y{:}X, A$, it will be understood that A is not in the set X. We shall usually abbreviate $X{:}\{A\}$ as $X{:}A$. The notation $X{:}[A]$ will stand for either $X{:}A$ or $X{:}\emptyset$. The notation $X,[A]{:}[B]$ will stand for either $X,A{:}B$ or $X{:}B$ or $X,A{:}\emptyset$. Within a sequent, commas will be used to represent set-unions. Thus X, Y will mean $X \cup Y$; and X, A will mean $X \cup \{A\}$.

Definition: $X_1 : Y_1$ is a *subsequent* of $X_2 : Y_2$ if and only if $X_1 \subseteq X_2$ and $Y_1 \subseteq Y_2$; and is a *proper* subsequent just in case at least one of these containments is proper. When $X_1 : Y_1$ is a proper subsequent of $X_2 : Y_2$ we also say that $X_2 : Y_2$ is a *weakening* of $X_1 : Y_1$, and that $X_1 : Y_1$ is a *strengthening* of $X_2 : Y_2$.

At the request of a referee, we have opted to use the term 'weakening' here rather than its synonyms 'dilution' or 'thinning'. Thus the more elegant phrase 'dilution-free', in the author's earlier publications, here becomes the rather awkward phrase 'weakening-free'. This stylistic loss will be offset, however, by the consideration that we shall be needing quite frequently to refer to the converse—the strengthening of a sequent. (In [20], stengthening was called 'concentration', the apposite antonym of 'dilution'.)

Definition: When a natural deduction Π has conclusion A and its undischarged assumptions form the set X, we say that the sequent $X : A$ is the *result* established by Π. A natural deduction Π_1 is *stronger* (resp. *weaker*) than another natural deduction Π_2 if and only if the result established by Π_1 is *stronger* (resp. *weaker*) than the result established by Π_2. We shall say that Π_1 is *at least as strong as* Π_2 when it is not the case that Π_2 is stronger than Π_1. In this case Π_1 will establish the same result as Π_2, or a stronger one.

1.3 Remarks on general method: historical summary

This investigation concerns only intuitionistic logic I, with an eye to extracting a certain subsystem of it, namely intuitionistic relevant logic *IR*, in a particularly natural way.

We exploit an approach to the elimination rules first bruited by Schroeder-Heister [8]. On this approach the elimination rules for $\wedge$ and $\rightarrow$ are, as we prefer to say, 'parallelized'. (A less informative term that has seen some use in this regard is 'generalized'.) The parallelized forms of these rules have been employed independently also by Paulson ([5], p. 23), as what he called 'destruction rules'.

It is important to appreciate that the two relevant systems *IR* and *CR* arise from relevantizing the *deducibility* relations of I and C respectively, and do not seek to produce a stricter, 'intensional' implication connective $\rightarrow$ in the respective object-languages. The main departure from logical orthodoxy in the systems *CR* and *IR* is that they give up the unrestricted transitivity of a Tarskian (also sometimes called a Scott-type) consequence relation. As we shall see below, however, we are fully compensated for this 'loss' of transitivity of deduction by the *epistemic gain* that thereby accrues. The 'restricted' transitivity that is afforded by these relevant systems is still adequate for all scientific purposes—both for mathematics and for the hypothetico-deductive method in natural science. (See below.)

Our preferred method of relevantizing is proof-theoretic; it was first introduced, in order to obtain *classical* relevant logic *CR* from C, in [9] and [10] for natural deduction, and in [11] for the sequent calculus. That *CR*, with its eschewal of unrestricted transitivity of deduction, is what is nowadays known as a 'substructural' logic is clear from the presentation in [11]. (Wansing [28] cites only [21] in this regard; but [11] serves just as well.)

The *intuitionistic* system *IR* can be extracted from I by the same method of relevantizing that produces *CR* from C. Systems of rules for both natural deduction and the sequent calculus were provided for *IR* (without $\rightarrow$) in [14]; and for the full languages of *IR* and *CR* in *Autologic*; and [23], ch. 10. Restall ([7], p. 150) describes the restricted transitivity of *IR* as 'idiosyncratic'—which at least licenses the inference to the conclusion that the suggestion that one restrict transitivity of deduction in this way is *original*. As will be argued below, however, such restriction is an effective prophylactic against irrelevance, and involves no logical losses inimical to scientific purposes. Indeed, the restriction actually leads to specifications of better proof-search algorithms for the unrestricted systems (see *Autologic*).

In the *natural deduction* setting, the method of relevantizing involves: banning the absurdity rule (ex falso quodlibet); banning vacuous discharge of assumptions in certain indirect rules, such as ¬-Introduction; and requiring proofs to be in normal form. In the *sequent* setting, the method of relevantizing is to have no 'structural' rules other than the rule of initial sequents $A:A$; and to frame the rules for the logical operators so as to ensure non-vacuous discharge of assumptions where necessary. Further details can be found in *Autologic*; and in [23], ch. 10.

It is important to realize that in the systems of natural deduction for *IR* and for *CR*, discharge-rules always discharge *all* occurrences of the assumption(s) in question; and the same effect is secured, in the sequent setting, by treating sequents as having *sets of sentences* as antecedents and succedents, rather than *sequences of sentence-occurrences*. Thus the whole issue of 'contraction' is orthogonal to our main concerns about relevance. The structural rule of contraction features only when antecedents of sequents are taken to be sequences of sentence-occurrences rather than sets of sentences. (Similar remarks apply to permutation and expansion.)

1.4 Features of IR*: historical summary*

The adequacy of first-order *IR* for the hypothetico-deductive method in natural science was demonstrated in [15], by simple adaptation of the considerations concerning M that were first entered in [12]; and the issue if *IR*'s adequacy for natural science was re-capitulated in [23].

The adequacy of intuitionistic relevant logic for intuitionistic mathematics was demonstrated in [19]. This is an immediate consequence of the Conservation Theorem stated below, which holds also at first order (cf. [19]).

Philosophical and meaning-theoretic arguments for the choice of *IR* as the logic most faithful to the anti-realistically licit meanings of the logical operators were set out in [13], [15] and [23]).

An exhaustive comparison of the relationship between *IR* and all other major systems of logic of a relevantist or intuitionistic character was provided in [16] and [22]. These other systems include M, I, Anderson-Belnap R, Meyer's *LR*, Došen's intuitionistic fragment R_{int} of R, the author's system of 'truth-table logic', and Kalmar logic.

Systematic proof-theoretic investigations of M and of *IR*, and the application of special normal-form theorems to provide very effective constraints on automated proof-search in those systems, were set out in *Autologic* and in [18]. The decision problem for propositional theoremhood in *IR*, like that of its parent system I, is PSPACE-complete. Moreover, the decision problem for propositional theoremhood in *CR*, like that of its parent system C, is co-NP-complete. Thus our method of relevantizing, unlike that of Anderson and Belnap's system *R*—which is *undecidable* (Urquhart [25])—does not make it harder to find proofs of a given conclusion from a given set of premisses. On the contrary; extensive experience with implementations of proof-search algorithms for *IR*—algorithms which exploit the relation of relevance that *IR* secures between premisses and conclusions—shows that this analysis of deductive relevance really does yield highly efficient heuristics that help to keep proof-search focused, without any loss of logical completeness. For a metatheorem furnishing a strong relevance constraint among premisses and conclusion of any provable sequent, see *Autologic*, ch. 9. It would of course be impossible, in either the classical or the intuitionistic case, for this method of relevantizing to *reduce* the complexity of the decision problem in the propositional case, since the relevant subsystem contains all theorems of the parent system (see below). Where the relevant subsystem differs from the parent system is in the deducibility relation between (finite, non-empty) sets of premisses and the conclusions to be drawn from them.

The epistemological and computational reflections required to justify the 'loss' of unrestricted transitivity of deduction within *CR* and *IR* were undertaken in [21], and [23], ch. 10. The following subsection summarizes these.

1.5 Restricted transitivity of deduction and epistemic gain

The systems *CR* and *IR* do not obey the unrestricted principle of transitivity. In the sequent setting for the classical system C, this principle is expressed by the structural rule of *Cut* in the following form:

$$\frac{X:Y,A \qquad A,Z:W}{X,Z:Y,W} \qquad \text{(where } A \text{ does not occur in either } Y \text{ or } Z\text{)}$$

whereas for the intuitionistic system I, the rule of Cut would take the form

$$\frac{X:A \qquad A,Z:B}{X,Z:B} \qquad \text{(where } A \text{ does not occur in } Z\text{)}$$

By contrast with the Cut rules for C and I, one can only say the following about *CR* and *IR* respectively, which lack any explicit rule of Cut:

(CR_1): if there are *CR*-proofs of the sequents $X:Y,A$ and $A,Z:W$, then there is a *CR*-proof of some (possibly *proper*) subsequent of $X,Z:Y,W$; and

(IR_1): if there are *IR*-proofs of the sequents $X:A$ and $A,Z:B$, then there is an *IR*-proof of some (possibly *proper*) subsequent of $X,Z:B$.

Indeed, these claims follow, respectively, from the following stronger claims:

(CR_2): from any *CR*-proofs of the sequents $X:Y,A$ and $A,Z:W$, one can effectively determine a *CR*-proof of some (possibly *proper*) subsequent of $X,Z:Y,W$; and

(IR_2): from any *IR*-proofs of the sequents $X:A$ and $A,Z:B$, one can effectively determine an *IR*-proof of some (possibly *proper*) subsequent of $X,Z:B$.

We can therefore say that the 'loss' of transitivity involved in giving up the orthodox principle of *absolutely unrestricted* transitivity of deduction (in the parent system C or I) is loss that *ought* to be incurred; for it is offset by the *epistemic gain* that accrues from obtaining a *stronger* logical result (within the respective relevantized subsystem *CR* or *IR*) than the 'lost' one. That a stronger logical result is always to be had when unrestricted transitivity fails is obvious from the results just stated. By way of illustration in the intuitionistic case, consider the following easy corollary of (IR_2):

> Suppose that one has *IR*-proofs of the sequents $X : A$ and $A, Z : B$, but that there is no *IR*-proof of the sequent $X, Z : B$. Then one can effectively find an *IR*-proof of a sequent of the form $\emptyset : B$, or of the form $V : B$ or $V : \emptyset$, for some non-empty *proper* subset V of X, Z. That is, we discover either that
> (1) B is a logical theorem (hence needs no support from any members of X, Z); or that
> (2) B follows from some of the available premisses in X, Z, but not all of the latter are needed; or that
> (3) some of the premisses in X, Z are jointly inconsistent.

Each of the alternative discoveries (1), (2) and (3) involves epistemic gain, in the form of a stronger logical result; so there is no point in mourning the loss of unrestricted transitivity of deduction. It fails where it *ought* to fail. A decisive illustration of such welcome failure, for the relevantist, is to be had from Lewis's proof of his first paradox:

$A : A \vee B$ (by $\vee$-Introduction);
$A \vee B, \neg A : B$ (Disjunctive Syllogism); whence, by Cut,
$A, \neg A : B$.

The system *IR* furnishes the first two deducibility statements. But it furnishes no proof for the sequent $A, \neg A : B$. Instead, it furnishes a proof only of the stronger logical result $A, \neg A : \emptyset$. Why 'stronger'?—because $A, \neg A : \emptyset$ is a *proper* subsequent of $A, \neg A : B$. With this degenerate example, the inconsistency of the premisses strikes the eye immediately. But this will not happen in general, when there are several complex premisses on the left. Hence the 'subsetting down' to $\emptyset$ on the right will often be a surprising and welcome discovery. Similar remarks apply to other patterns of 'subsetting down' in the general case.

Note that the failure of unrestricted transitivity of deduction has no effect whatsoever on the stock of *theorems*. Indeed, any theorem of C is a theorem of *CR*; and any theorem of I is a theorem of *IR*. Thus, given the completeness of C, we can say that any classical tautology is a theorem of *CR*. Likewise, given the completeness of I for intuitionistically valid sentences (i.e. sentences true in all Kripke models), we can say that any intuitionistic validity is a theorem of *IR*.

These statements about theoremhood are special cases of a more general kind of completeness of the relevant subsystems *CR* and *IR*.

Definition: A sequent is *perfectly valid* in C (resp. I) just in case it is valid in C (resp. I) but has no proper subsequent that is valid in C (resp. I).

Theorem. Every perfectly valid sequent in C (resp. I) is provable in *CR* (resp. *IR*).
Proof. Every classical proof Π of A from X can be turned into a classical relevant proof Π' of A or of $\perp$ from some subset X' of X (see [11]). But if $X : A$ is perfectly valid in C, then X' must be X itself, and the conclusion of Π' must be A, not $\perp$.

Similarly, every intuitionistic proof Π of A from X can be turned into an intuitionistic relevant proof Π' of A or of $\perp$ from some subset X' of X (see [15], [19]). But if $X : A$ is perfectly valid in I, then X' must be X itself, and the conclusion of Π' must be A, not $\perp$. *QED*

How can *CR* (resp. *IR*) be complete for theorems of C (resp. I) and yet differ on deducibilities? It is because the usual *deduction theorem* technically fails, in one direction, for these relevantized systems. But that failure turns out to be a boon! Again, we illustrate with reference to the intuitionistic case. In one direction, the deduction theorem holds:

> if there is an *IR*-proof of B that uses (exactly) all the premisses in X, then there is an *IR*-proof of $A \rightarrow B$ using exactly the premisses in $X \setminus \{A\}$.

In the converse direction, however, the deduction theorem fails. For:

> there is an *IR*-proof of the sequent $\neg A : A \rightarrow B$, while there is no *IR*-proof of the sequent $A, \neg A : B$.

This serves to underscore the point made above: we are concerned to relevantize the *deducibility* relation rather than the object-language conditional.

1.6 Non-forfeiture of epistemic gain in automated deduction

The restricted principle of transitivity of deduction available in the systems *CR* and *IR* leads to further benefits for the computational logician seeking to automate proof-search. Both these logical systems obey a principle that was called, in [23], the *Principle of Non-Forfeiture of Epistemic Gain.* This principle can be explained as follows.

In automated proof-search one is constantly pursuing sub-goals of the general form 'Find a proof of the sequent $X : Y$'. And often it transpires that such sub-goals succeed spectacularly, with the discovery of a proof of some 'very proper' subsequent of $X : Y$. In various other systems of relevant logic, however, these stronger results might turn out to be to no avail, because of the need to have *used* various assumptions for subsequent discharge lower down in the proof-tree that is being built up from below. In these other systems of relevant logic, the requirements on how those assumptions should have been used is artificially exigent. (See [6], for example, for the restrictions on $\rightarrow$-Introduction in the Anderson-Belnap system R.) Thus those spectacular successes higher up in the evolving proof-tree will very often have to be discarded, and alternative sub-proofs sought. This is not, however, the case with *CR* and *IR*. Because of the way these systems have been 'sensibly' relevantized, one can *always* exploit the discovery of a proof of a stronger-than-expected result for a subgoal, by incorporating the proof of the stronger result at the appropriate point in the unfolding of the search- (and hence proof-) tree. In general the discovery of a stronger-than-expected result in pursuit of a subgoal will contribute to the discovery of a proof of a similarly stronger-than-expected overall result—that is, a proof of a proper subsequent of the sequent involved in the original goal. This is the deep underlying reason why the notion of relevance uniformly captured in *CR* and in *IR* makes heuristics for proof-search in those systems (hence, in their parent systems C and I) so much more efficient.

2 A summary of results in and about *IR*

Definition: When we say that a sequent $X:A$ is provable in *IR* (or that $X:A$ is an *IR*-deducibility) we mean that there is a natural deduction in *IR* whose conclusion is A and whose premisses (undischarged assumptions) form *exactly* the set X; or that there is a proof in the sequent system for *IR* whose bottom sequent is *exactly* $X:A$. Note that any sequent of the form $X:\emptyset$ can be rendered, in the natural deduction setting, as $X:\bot$, where $\bot$ is the absurdity symbol.

In order to help orient the reader with little or no familiarity with the system *IR*, we provide the following list of deducibilities and non-deducibilities in *IR*. In due course the full set of inference rules will be stated, once they have been made to emerge naturally from the discussion. The two lists below serve, at this stage, to provide some useful orientation.

Deducibilities of *IR*	Non-deducibilities of *IR*
$A, \neg A:\emptyset$	$A, \neg A:B$
$A:\neg\neg A$	$A, \neg A:\neg B$
$A, B:A \wedge B$	
$A \wedge B:A$	
$A \wedge B:B$	
$A:A \vee B$	
$B:A \vee B$	
$A \vee B, \neg A:B$	
$A \vee B, \neg A, \neg B:\bot$	
$B:A \rightarrow B$	
$\neg A:A \rightarrow B$	
$A \rightarrow B, A, \neg B:\bot$	
$A \wedge \neg A:A$	
$A:A \wedge A$	
$A \wedge \neg A:\neg A$	
$A:\neg A \rightarrow B$	
$\emptyset:A \rightarrow (\neg A \rightarrow B)$	
$\emptyset:A \rightarrow (A \rightarrow A)$	
$\emptyset:A \rightarrow (B \rightarrow A)$	
$\emptyset:(A \rightarrow \neg A) \rightarrow \neg A$	

Any deducibility with empty antecedent (i.e., a theorem) is in *IR* because it is in I. Note that the left-hand list contains every deducibility statement that can be 'read off' any line in the truth-tables for the connectives. Note also that it contains disjunctive syllogism. This means that *IR* cannot be contained in the intuitionistic fragment of the Anderson-Belnap system R, since the latter does not have disjunctive syllogism; nor can *IR* be contained in Johansson's *Minimalkalkül* M, since the latter also lacks disjunctive syllogism. Note, finally, that *IR* lacks $A, \neg A : \neg B$. This means that *IR* does not contain M.

The following are the most important metatheorems about *IR*. (We omit the analogues for *CR*.)

- *Conservation Theorem*: Any normal natural deduction in I of A from X can be transformed, in linear time, into a natural deduction in *IR* of either A or $\perp$ from some subset of X.
- *Corollary* (given the completeness and normalization theorems for I): If $X : A$ is intuitionistically valid, then some subsequent of $X : A$ is provable in *IR*.
- *Further corollaries*:
  1. Every perfectly intuitionistically valid sequent is provable in *IR*;
  2. Every theorem of I is a theorem of *IR*;
  3. If X is intuitionistically inconsistent, then there is an *IR*-proof of the inconsistency of some subset of X.
- The problem of deciding whether a finite sequent of propositional logic has an *IR*-provable subsequent is PSPACE-complete.
- *Principle of Restricted Transitivity of Deduction, and Epistemic Gain*: Given any *IR*-proofs of the sequents $X : A$ and $A, Z : B$, one can effectively find an *IR*-proof of some (possibly *proper*) subsequent of $X, Z : B$.

3 Towards a fruitful analysis of relevance

Given what is already known about *IR*, the question arises whether there is any further light that proof-theoretic investigations might shed on the system. It turns out that there is. In *Autologic*, the conventional Gentzen-Prawitz-style natural deduction system was replaced by a form of natural deduction that was called a *hybrid* system. Proofs in the hybrid system look like (indeed, are) natural deductions, in that their nodes are labelled by single sentences (not sequents), and their branchings are applications of I-rules and E-rules, some of which can discharge assumptions at leaf-nodes. All E-rules, moreover, are parallelized. But a hybrid proof's overall tree-structure is that of a cut-free sequent-proof. Thus the hybrid system is a combination of the natural deduction system and the Gentzen sequent system, seeking to incorporate the advantageous features of each while avoiding the disadvantageous ones. The most important feature of the hybrid rules, as stated in *Autologic* (ch. 5, p. 41) was that

> Major premisses for elimination 'stand proud' in our proof trees; they never stand as conclusions of any rules. This makes our formulation a *hybrid* between the Prawitz-style natural deduction formulation and the Gentzen-style sequent formulation. It also means, in effect, that we allow *only proofs in normal form.* This is specially advantageous in constraining proof-search.

The point was reiterated a few pages later (op. cit., ch. 6, p. 47 and p. 50) as follows:

> Not every Prawitz proof is normal. A proof in the hybrid system (using rules in parallel form), however, is normal *by definition*, in that major premisses for eliminations have to 'stand proud'. They therefore do not stand as conclusions of introductions. *Nor do they stand as conclusions of eliminations.* And this

> is what limits one's freedom to re-order steps in a hybrid proof. ... A hybrid proof encodes the structural information of a sequent proof, but does so more economically. [Emphases in the original.]

It is clear from the context that the sequent proof in question would always be cut-free.

3.1 Historical digression: an issue of priority

Autologic's move, back in 1992, of having all major premisses for eliminations stand proud, or, equivalently, stand as *assumptions* within the proof, is adopted also by von Plato ([26], [27]), and by Negri and von Plato ([4]), but without attribution.

At p. 123 of [26] we read 'All major premisses of elimination rules in the derivation are assumptions, which is the characteristic property of normal derivations with general elimination rules.' Compare also [4], at p. xvii: 'A derivation [in natural deduction] is normal when all major premisses of elimination rules are assumptions.' From the omission of any citation of *Autologic* regarding this crucial point, it would appear that von Plato and Negri must have been unaware of *Autologic*'s much earlier use of this move.

The matter of normal form for the system of parallelized rules is actually a good bit more delicate than might at first appear. As we shall see below, the Prawitzian notion of normal form in the setting of serial rules can be straightforwardly transferred to the setting of parallelized rules. But the transferred notion does not yet carry the requirement that major premisses for eliminations should stand proud (as assumptions). *That* requirement goes over and above the normal notion of normality, so to speak. The requirement can indeed be imposed, to obtain a more exigent notion of normal form, which we shall call *exposed* normal form. (This is essentially what von Plato calls 'full normal form', if we turn a blind eye to his allowing $\perp$ to stand as an assumption.) But then further non-trivial justification is required for the claim that every proof can be converted to one in this more exigent normal form.

In [27], the reader does indeed find such justification, thereby learning that von Plato reprises, with inessential variations but in a very roundabout manner (see below), the heart of the treatment of hybrid proofs given nine years earlier in *Autologic*. See *Autologic*, p. 41, for a full anticipation of the 'key insight' (that one should parallelize the elimination rules) which von Plato claims in his 2001 paper, in the Introduction on p. 541. The crucial feature (see Definition 3.3, p. 549) of his notion of 'full normal form' for a natural deduction is the idea from *Autologic*, that all major premisses for eliminations (MPEs) are assumptions. The main result of [27] is stated as follows (p. 559):

> Given a non-normal derivation [in natural deduction], translation to sequent calculus, followed by cut-elimination and translation back to natural deduction, will produce a [full] normal derivation:
>
> **Theorem 5.1. Normalization.** Given a natural deduction derivation of C from Γ, the derivation converts to a normal derivation of C from Γ^* where each formula in Γ^* is a formula in Γ.

Again, in [4], the main result of Chapter 8 is stated as follows:

> A translation from non-normal derivations [in the system of natural deduction] to derivations [in the sequent salculus] with cuts is given, from which follows a

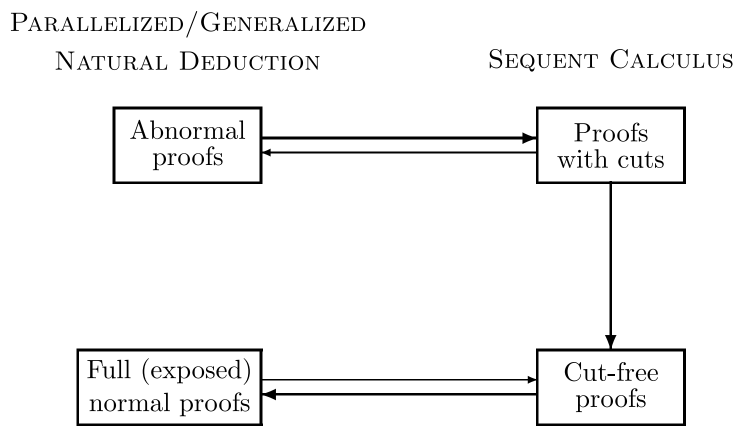
> normalization procedure [for natural deductions] consisting of said translation, followed by cut elimination and translation back to natural deduction.

The neglect of *Autologic* by Negri and von Plato is complete when one encounters the following claim, on p. 208 of their book:

> No one seems to have followed the idea that it is this system of natural deduction [i.e., the system 'with the usual &E rules'], not sequent calculus, that lies at the back of the failure of isomorphism between derivations in the two calculi.

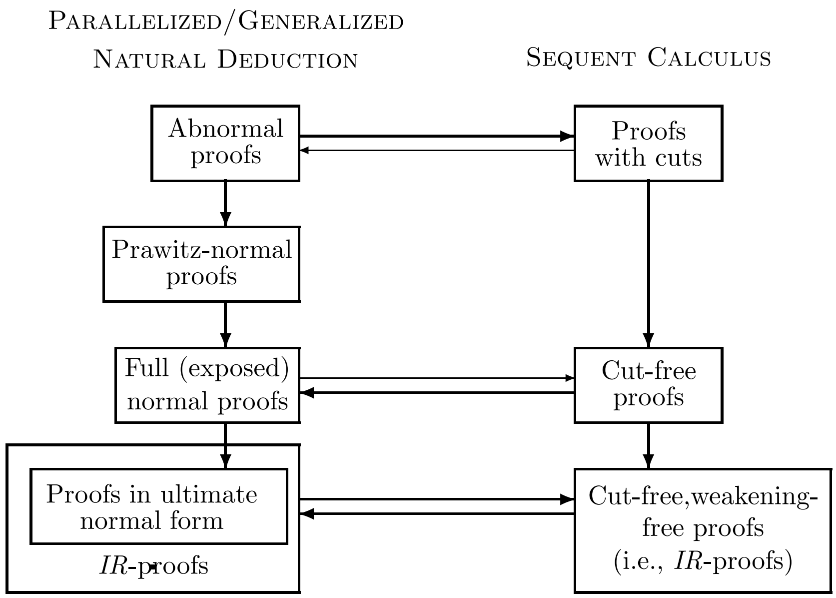
It was one of the central projects of *Autologic* to overcome that failure of isomorphism, precisely by using its so-called 'hybrid' proofs, which are none other than the proofs with parallelized elimination rules that Negri and von Plato discuss, and for which (full) normality is definable as consisting in all major premisses standing as assumptions.

Despite the fact that [27] and [4] reprise these leading ideas from *Autologic*, the resulting situation, on their above-quoted accounts of it, is conceptually unsatisfying. In effect, [27] and [4] negotiate the five arrows in the following diagram, three of which—indicated with thicker lines in the diagram—are important for present purposes. What these authors show is merely that beginning with a natural deduction not in normal form, one can travel east, then south, then west, to arrive at a natural deduction in full (exposed) normal form. The lateral arrows register rather trivial correspondences. Only the vertical arrow registers a non-trivial result, which of course can be taken directly from Gentzen [1].



Why is this conceptually unsatisfying? For the reason that *we seek an analysis of the notion of normal form for natural deductions, suitably formalized, that will allow one to establish Gentzen's* Hauptsatz *directly for the system of natural deduction, without the need for such a meta-detour through the sequent calculus.* Remember that a 'Hauptsatz' is a result to the effect that one can always find proofs without detours, whichever system of proof one might be using. (Thus the Cut Elimination Theorem is *a* Hauptsatz, in that it is the Hauptsatz for the sequent calculus; it is not *the* Hauptsatz *überhaupt.*) Gentzen was driven to consider the sequent calculus precisely because the Hauptsatz for *natural deduction systems* eluded him. So the real conceptual and theoretical advance that is worth pursuing here is one that would allow us to arrive at 'full' (or exposed) normal forms *directly*, using only the structural materials afforded by the natural deduction systems themselves.

This indeed is what we shall accomplish in this paper. We shall see that there is, for the parallelized system of natural deduction, an intermediate notion of normal form—let us call it *Prawitzian normal form*—that can be interpolated between abnormal forms and full (or exposed) normal forms. Moreover, using only procedures applying to natural deductions themselves, and without venturing to consider sequent proofs (let alone trying to prove cut-elimination!), we shall show that any natural deduction can be converted, first, into Prawitzian normal form, and thence into full (or exposed) normal form. Finally, we shall go one important step further, and convert exposed normal forms into *ultimate* normal forms, which will afford our sought analysis of relevance:



López-Escobar ([3]) is another proof-theorist apparently unaware of the system of hybrid proof in *Autologic*. Although both von Plato and López-Escobar work with parallelized (generalized) elimination rules, neither of them seizes the opportunity afforded by the new kinds of normal form thereby available to give a fruitful analysis of *relevance*. Indeed, both these writers not only retain the absurdity rule—the relevantist's nightmare—but force it into an unnaturally complicated mould in order to have all 'elimination' rules (among which they reckon the absurdity rule) take the same form. They allow $\perp$ to feature as a sentential constant, even allowing it to appear as an assumption. By contrast, the approach favoured here treats $\perp$ as at most a punctuation marker within proofs, whose function is only to register a contradiction. For principled arguments for so confining the role of $\perp$, see [24].

4 Ultimate normal forms

The interesting and difficult question is whether a translation of cut-free sequent proofs into normal natural deductions, and its converse, can be exploited in a fruitful analysis of *relevance* of premisses to conclusions. These investigations are devoted to providing an answer in the affirmative, showing how naturally the system *IR* arises from such an analysis. But in order to obtain this analysis, we have to set up the correspondence between, on the one hand, natural deductions in *ultimate normal form*, and, on the other hand, *cut-free, weakening-free* sequent proofs. We indicated this by means of the lowest lateral arrows in the last diagram above.

Our aim is to take the method of normalization to a natural extreme. We begin with the system I of intuitionistic logic, and work towards the relevantized subsystem *IR*. We define an extremely exigent notion of normal form, and show that all intuitionistic proofs can be converted into proofs of this form. The 'ultimately' normalized version Π' of an intuitionistic proof Π of a conclusion A from a set X of undischarged assumptions will itself have as its conclusion either A or $\perp$ (absurdity), and will have its undischarged assumptions drawn from the set X. Thus in an obvious sense the normalized proof Π' will establish a logical result that is at least as strong as, and possibly stronger than, that established by the non-normalized proof Π. The possibility of stronger logical results arises from the way in which we shall have got rid of all applications of the absurdity rule that might have occurred in Π.

There is another respect in which the notion of ultimate normal form employed herein is more demanding than usual. In proofs in ultimate normal form, because of the sorts of transformation that have to be applied in order to produce them from non-normal proofs, every major premiss for an elimination will 'stand proud'; that is, every MPE will occupy a leaf-node of the proof-tree. This feature allows one to set up the aforementioned natural isomorphism between any natural deduction in ultimate normal form and the corresponding sequent proof. Indeed, the sequent proof just *is* the natural deduction, once one prunes off those leaf nodes occupied by MPEs, and re-labels every node of the natural deduction with the sequent (rather than the conclusion) that has been established by that stage.

This is somewhat surprising, given how in Gentzen's treatment there was quite a considerable gap between sequent proofs and even those natural deductions that were normal in his sense. The gap is narrowed by following *Autologic*'s strategy of framing all elimination rules in 'parallel' form. Gentzen did not use these parallel forms. Instead, he chose 'serial' forms for (→E) and (∧E), thereby estranging natural deductions from their corresponding sequent proofs.

In standard treatments, transformation to normal form is effected by applying the so-called *reduction procedures*, whose aim is to rid a proof of any sentence-occurrence standing as the conclusion of an introduction (or of an application of the absurdity rule—see below) and as the major premiss of the corresponding elimination. Here, we go much further. We employ further transformations (to be explained presently) with suggestive names like 'shuffling', 'exposing' and 'extracting', in order to complete the process of *ultimate* normalization. The net effect of these transformations is to *relevantize* the proof. *Every intuitionistic proof gets transformed into an intuitionistic relevant proof.* Intuitionistic relevant logic is thereby distinguished as the system in which this transformational analysis of the essence of the line of reasoning 'within' an intuitionistic proof is laid bare, and in such a way as to reveal the resulting normal natural deductions and their corresponding sequent proofs to be one and the same structure.

5 Rules of natural deduction for M and I

The system M of minimal propositional logic consists of the following rules for introducing and eliminating the connectives. (The boxes and diamonds will be explained in the next subsection.)

	INTRODUCTION	ELIMINATION
$\neg$	$\dfrac{\begin{array}{c}\Diamond\,\overline{A}^{(i)}\\ \vdots \\ \bot\end{array}}{\neg A}{(i)}$	$\dfrac{\neg A \quad A}{\bot}$
$\wedge$	$\dfrac{A \quad B}{A \wedge B}$	$\dfrac{A \wedge B \quad \begin{array}{c}{}^{(i)}\overline{A}\ \Box\ \overline{B}^{(i)}\\ \underbrace{A\ ,\ B}\\ \vdots \\ C\end{array}}{C}{(i)}$
$\vee$	$\dfrac{A}{A \vee B} \qquad \dfrac{B}{A \vee B}$	$\dfrac{A \vee B \quad \begin{array}{c}\Box\,\overline{A}^{(i)}\\ \vdots \\ C\end{array} \quad \begin{array}{c}\Box\,\overline{B}^{(i)}\\ \vdots \\ C\end{array}}{C}{(i)}$
$\rightarrow$	$\dfrac{\begin{array}{c}\Diamond\,\overline{A}^{(i)}\\ \vdots \\ B\end{array}}{A \rightarrow B}{(i)}$	$\dfrac{A \rightarrow B \quad A \quad \begin{array}{c}\Box\,\overline{B}^{(i)}\\ \vdots \\ C\end{array}}{C}{(i)}$

Definition: In applications of the rule of $\rightarrow E$, the sub-proof with conclusion A is called the *minor* sub-proof, while the sub-proof within which the assumption B is discharged is called the *major* sub-proof.

Note that there are no constraints, other than those indicated with boxes and diamonds, on how the rules may be applied. One is free to use them in any order. Hence one can construct proofs that are not in normal form. An abnormal proof might contain, for example, sentence-occurrences that stand as conclusions of I-rules and as the major premisses of applications of the corresponding E-rules. Such occurrences are called 'maximal'. More generally, abnormal proofs might contain maximal *chains* of occurrences of some same sentence. For the notions of chain and of maximal chain, see below.

The system I of intuitionistic logic is obtained from minimal logic simply by appending the *absurdity rule*, or *ex falso quodlibet*:

$$\dfrac{\bot}{A}$$

An example of a result that can be proved in intuitionistic but not in minimal logic is *disjunctive syllogism*, $A \vee B, \neg A : B$. The proof is

$$\dfrac{A \vee B \qquad \dfrac{\dfrac{\neg A \qquad \overline{A}^{\,(1)}}{\bot}}{B} \qquad \overline{B}^{\,(1)}}{B}\,(1)$$

Once again, we stress (as we did for M) that proofs in I need not be in normal form. Indeed, I allows a new kind of 'maximal sentence occurrence': one standing as the conclusion of an application of the absurdity rule, and as the major premiss of an elimination.

Definition: A *maximal sentence-occurrence* in an intuitionistic proof is one that stands as the conclusion of an introduction or of an application of the absurdity rule, and as the major premiss of an elimination.

Note that if a maximal sentence-occurrence is the conclusion of an introduction, then that occurrence will be the major premiss of the corresponding elimination rule. This is because introductions introduce, and eliminations eliminate, *dominant* occurrences of logical operators. A maximal sentence occurrence represents an unnecessary detour; it embodies, locally, some needless logical complexity—that of its dominant operator.

5.1 Vacuous v. non-vacuous discharge of assumptions

Note that minimal logic (and therefore also intuitionistic logic) is quite permissive with regard to the *discharges of assumptions* allowed by certain of the rules given above. The permissive construal is that one is permitted to discharge any assumption of the indicated form, *if* one has used it. The permissive construal allows one to apply a so-called 'discharge'-rule even in cases where no assumption of the indicated form has been used. In such a case one could speak of 'vacuous discharge'. In our statement of the rules above, this permissive construal applies to the rules $(\neg I)$ and $(\rightarrow I)$. Here, permissiveness is indicated by a diamond appended to the discharge-stroke. For all the other discharge-rules, we adopt instead a construal according to which discharge is *obligatory*. Obligatoriness is indicated by a box appended to the discharge-stroke. Here, the affected rules may be applied *only if* at least one assumption of the indicated form(s) has indeed been used to obtain the sub-conclusion concerned. (For the parallel version of the rule of $(\wedge E)$ given above, one must have used either A or B as an assumption, but not necessarily both.)

There is in the literature on natural deduction a precedent for such insistence on obligatory discharges. Prawitz, for example, when speaking of normal forms of natural deductions in intuitionistic logic, explicitly rules out what he calls *redundant* applications of $(\vee E)$ and of $(\exists E)$. These are precisely applications of these rules that do not discharge any assumptions of the indicated form. (See [6], at pp. 49–50.) As Prawitz points out, such applications are superfluous. Remarks in this spirit would apply also to applications of our new parallel forms of $(\wedge E)$ and of $(\rightarrow E)$. (Prawitz used only the serial forms of these rules; see below.)

Note that the occasional insistence, within the formulation of the natural deduction rules just given, of a box on a discharge stroke—indicating the need for an assumption of the indicated form—is compatible with those rules' being used to construct proofs that are not in normal form. The boxes occur in the parallelized elimination rules in order to secure two-way translatability between proofs using the serial elimination rules and proofs using the parallelized forms of those elimination rules. In particular, we want to ensure that any proof constructed by means of the parallelized rules can be translated into a proof that uses the corresponding serial rules. We turn to this topic in the next subsection.

5.2 Serial v. parallel forms of elimination rules

Note how all the elimination rules, including those for conjunction $(\wedge)$ and implication $(\rightarrow)$, are in *parallel* form. The usual 'serial' form of $(\rightarrow E)$ (so-called *modus ponens*) is

$$\frac{A \quad A \rightarrow B}{B}$$

This is the form of $(\rightarrow E)$ adopted by both Gentzen and Prawitz. (See the English translation of [1], at p. 77. See also [6], p. 20.) To see that this is equivalent to our 'parallel' form of $(\rightarrow E)$, imagine that the serial form has been used in a proof as follows:

$$\begin{array}{c} \begin{array}{cc} & \begin{array}{cc} X & Y \\ \Pi & \Sigma \\ A & A \rightarrow B \end{array} \\ \end{array} \\ \underbrace{Z \qquad , \qquad \overline{\qquad\quad B \qquad\quad}}\ \text{serial}\ (\rightarrow E) \\ \Theta \\ D \end{array}$$

Then the component sub-proofs Π, Σ and Θ can be re-arranged as follows:

$$\dfrac{\begin{array}{c}Y\\ \Sigma\\ A\rightarrow B\end{array}\qquad \begin{array}{c}X\\ \Pi\\ A\end{array}\qquad \begin{array}{c}\underbrace{Z\ ,\ \overline{B}^{(i)}}\\ \Theta\\ D\end{array}}{D}{\scriptstyle (i)}\ \text{parallel}\ (\rightarrow E)$$

Note the potential economies in this re-arrangement. Consider the first proof-schema, in which the *serial* form of $(\rightarrow E)$ is used. If B enjoys several different assumption-occurrences within the sub-proof Θ, one would need each such assumption-occurrence to be 'covered' by the sub-proof ending with the step of serial $(\rightarrow E)$. So we would need that many applications of serial $(\rightarrow E)$. Thus there would be that many copies of the minor sub-proof Π and the major sub-proof Σ—a great deal of unnecessary replication.

Now consider by contrast the second proof-schema, in which the *parallel* form of $(\rightarrow E)$ is used. Even if B enjoys several assumption-occurrences within the sub-proof Θ, they are all taken care of by the *single* application of parallel $(\rightarrow E)$, which in one fell swoop *discharges* all the assumption-occurrences of B within Θ. The result is that *only one* copy of the minor sub-proof Π, and *only one* copy of the major sub-proof Σ need be used.

In this re-arrangement of steps so as to parallelize the application of $(\rightarrow E)$, we see the reason why nothing is lost by insisting on non-vacuous discharge with the parallel form of the rule. In the vacuous case (where B is not used as an assumption in Θ), there would have been no corresponding instance of serial $(\rightarrow E)$ to contend with. Indeed, non-vacuous discharge when applying the parallelized rule is required if one is to be able, in general, to re-construct the proof in a form that uses the corresponding serial rules.

The two patterns of overall proof are inter-transformable, but clearly the second pattern is more economical in general. That is why we prefer the parallel form of $(\rightarrow E)$, even though it may seem a little more complex than the serial form. That slight increase in apparent complexity secures huge gains in efficiency in searching for proofs and in setting them out once we have found them.

Similar considerations apply to the parallel v. serial form of $(\wedge E)$. Take any proof using the serial form of $(\wedge E)$ on the major premiss $A \wedge B$:

$$\begin{array}{c}\underbrace{Y\ ,\ \dfrac{\begin{array}{c}X\\ \Pi\\ A\wedge B\end{array}}{A}\ ,\ \dfrac{\begin{array}{c}X\\ \Pi\\ A\wedge B\end{array}}{B}}\\ \Sigma\\ C\end{array}$$

This can be transformed into the following more economical proof-schema that uses the parallel form of $(\wedge E)$:

$$\dfrac{\begin{matrix} X \\ \Pi \\ A \wedge B \end{matrix} \qquad \begin{matrix} \underbrace{Y \ , \ \overline{A}^{(i)} \ , \ \overline{B}^{(i)}} \\ \Sigma \\ C \end{matrix}}{C}{\scriptstyle (i)}$$

Once again, only one copy of Π is needed here, rather than the possibly multiple copies that might be needed in the serial case.

As before with $(\rightarrow E)$, we see in the re-arrangment of steps so as to parallelize the application of $(\wedge E)$, the reason why nothing is lost by insisting on non-vacuous discharge with the parallel form of the rule. At least one of the conjuncts A, B should have been used as an assumption in the subordinate proof for parallel $(\wedge E)$. If this were not so, there would have been no corresponding instance of serial $(\wedge E)$ to contend with.

The other elimination rules (for $\neg$ and for $\vee$) are already in parallel form, and have no 'serial form' alternatives.

6 Finding ultimate normal forms of proofs in I

When we speak, loosely, of a 'system of parallelized rules' we mean, of course, the system of rules that includes the introduction rules, and all of whose elimination rules have been parallelized in the manner indicated above.

In order to understand the problem we have set ourselves, let us stress again that the system of parallelized rules that we have given above allows for proofs that are *not* in normal form. An anachronistic way of making the problem vivid is to ask the reader to imagine that when Gentzen devised his systems of natural deduction (his N systems), he had not opted for the serial forms of $\wedge$- and $\rightarrow$-elimination. Imagine that Gentzen had, instead, opted for Schroeder-Heister's parallelized forms of these rules. Our imaginary Gentzen might, nevertheless, have failed to achieve his *Hauptsatz*—that any deducibility can be established by a proof lacking detours—in the natural deduction setting, and consequently still have resorted to using his sequent systems to the same end, via the cut-elimination theorem. Then, when Prawitz later addressed the question of how to establish a *Hauptsatz* in the natural deduction setting, he would have been working with the above system of *parallelized* rules, inherited from our imaginary Gentzen. In order to establish a normalization theorem for this system, our imaginary Prawitz would have had to provide a different method of normalizing, with different outcomes. For, natural deductions constructed in the system of parallelized rules are different beasts, structurally, from ones constructed according to the serial rules.

Our investigations begin, then, with the problematic that would have confronted our imaginary Prawitz. We have this system of *parallelized* elimination rules, and we want to show that any proof constructed in accordance with them can be brought into normal form by a finite sequence of reduction steps. Given, however, how one defines 'normal form'—by faithful transfer of the Prawitzian notion in the serial setting—it will turn out that normal forms in the parallelized setting do not pack the punch that they could, without loss of completeness, be required to. We shall therefore be motivated to define a more exigent notion of normal form, and to show that the erstwhile normal forms can be brought into this more exigent form. We shall call the latter 'ultimate normal form'.

We shall now describe a series of different kinds of transformations on proofs, whose successive application will convert any intuitionistic proof into one in ultimate normal form. The kinds of transformations are:

(σ) shuffling $\vee$-eliminations downwards;

(ρ) reducing maximal chains (which will be defined presently);

(ϵ) exposing major premisses for eliminations; and

(*) extracting the relevant kernel of the proof (i.e., avoiding applications of the absurdity rule, and eschewing steps involving vacuous discharge of assumptions).

The 'Prawitz-style' normal form will have been achieved after the first two kinds of transformations (σ) and (ρ) have played themselves out (that is, shuffling $\vee$-eliminations downwards and then reducing maximal chains). The next kind of transformation (ϵ) exposes MPEs and thereby yields what we shall call the 'exposed normal form'. The final kind of transformation (*) then extracts the relevant kernel—thereby completing the transition to ultimate normal form.

6.1 *Shuffling down $\vee$-eliminations*

Any step of $\vee$-elimination can be 'shuffled downwards' over any immediately succeeding elimination step whose major premiss is the conclusion of that $\vee$-elimination, by means of transformations of the general form (cf. *Autologic*, p. 48):

$$\dfrac{\dfrac{\begin{array}{c}\Pi_1\\ A\vee B\end{array}\quad \begin{array}{c}\overline{\ \ }^{(i)}\\ A\\ \Pi_2\\ C\end{array}\quad \begin{array}{c}\overline{\ \ }^{(i)}\\ B\\ \Pi_3\\ C\end{array}}{C}{\scriptstyle (i)}\qquad (\Sigma_1)\ \ (\Sigma_2)}{D} \quad\overset{\sigma}{\longmapsto}\quad \dfrac{\begin{array}{c}\Pi_1\\ A\vee B\end{array}\quad \dfrac{\begin{array}{c}\overline{\ \ }^{(i)}\\ A\\ \Pi_2\\ C\end{array}\ (\Sigma_1)\ \ (\Sigma_2)}{D}\quad \dfrac{\begin{array}{c}\overline{\ \ }^{(i)}\\ B\\ \Pi_3\\ C\end{array}\ (\Sigma_1)\ \ (\Sigma_2)}{D}}{D}{\scriptstyle (i)}$$

Here the conclusion C of the $\vee$-elimination in the preimage is the major premiss of an elimination: precisely *which* elimination determining whether one has just one sub-proof Σ_1 to deal with, or both Σ_1 and Σ_2. With $\neg$-elimination and (parallel) $\wedge$-elimination there is just Σ_1; with $\vee$-elimination and (parallel) $\rightarrow$-elimination one has both Σ_1 and Σ_2.

Definition. Π is a *shuffled* proof if and only if no application of $\vee$-elimination within Π has its conclusion standing as the major premiss of an elimination.

Theorem. The process of shuffling $\vee$-eliminations downwards within a given proof Π must eventually terminate, yielding a (unique) proof $\sigma(\Pi)$ in which no major premiss for $\vee$-elimination is immediately succeeded by the major premiss of an elimination. The 'shuffled' proof $\sigma(\Pi)$ will obviously establish the same result as the original proof Π.

Proof. By repeated application of the transformation σ.

6.2 Converting shuffled proofs into normal form

We said above that a maximal sentence-occurrence embodies, locally, some needless logical complexity—that of its dominant operator.

Such unnecessary logical complexity can also arise when there is a *chain* of occurrences of some same sentence A within a proof.

Definition: A *chain* of length n is a descending sequence of n occurrences of some same sentence A (other than $\perp$), those occurrences of A lying consecutively on a branch within a proof, and with none of these occurrences of A standing as the conclusion of the minor sub-proof for a step of $\rightarrow$-elimination. We shall speak also of an A-*chain*. We shall call A the *link-sentence* of the chain. As such A is a sentence-type, which enjoys type-identical occurrences within the chain. A chain will be called *maximal* just in case (i) the first (topmost) occurrence of A within this chain stands as the conclusion of an introduction or of an application of the absurdity rule; and (ii) the last (bottommost) occurrence of A within this chain stands as the major premiss of an elimination.

Observation: It follows from this definition that any intermediate occurrence of A within a chain (whether or not it is maximal) is the conclusion of an elimination other than $\neg$-elimination.

Proof. An intermediate occurrence of A within a chain cannot be the conclusion of $\neg$-elimination, since by definition no chain can include any occurrences of $\perp$. Nor can such an occurrence of A be the conclusion of an introduction rule, since these rules introduce new complexity into their conclusions, whereas the chain must consist of occurrences of some same sentence. Nor can such an occurrence of A be the conclusion of the absurdity rule, since that would require the member of the chain immediately above that occurrence of A to be an occurrence of $\perp$, again contrary to the definition of chain. The Observation now follows.

The reader should bear in mind that, because $(\wedge E)$ and $(\rightarrow E)$ are being taken to be in *parallel* form, both these kinds of eliminations, as well as ($\vee$-E), can contribute to the propagation of a chain. This is because with eliminations of all three kinds, the main conclusion is of the same form as the conclusion of at least one of the subordinate proofs. Moreover, by the foregoing Observation, *only* these three kinds of step—$(\wedge E)$, $(\rightarrow E)$ and $(\vee E)$—can contribute to the propagation of a chain. Such an elimination contributes to the propagation of an A-chain by having one of the chain's occurrences of A as a minor premiss, and by having as its conclusion the immediately succeeding occurrence of A within the chain.

A maximal occurrence of a sentence A within a natural deduction is now a special case of a maximal A-chain, namely an A-chain of length 1, whose topmost occurrence of A is identical to its bottommost occurrence of A.

Observation. In a shuffled proof, no two distinct maximal A-chains can intersect.

Proof. Suppose for reductio that Π is a shuffled proof and that two distinct maximal A-chains intersect within Π. Then their point of intersection α must be an occurrence of A standing as the conclusion of $\vee$-elimination—since $\vee$-elimination is the only eligible elimination rule that involves more than one sub-proof alongside its major premiss, with conclusions equiform to the overall conclusion of the elimination. (Remember that by definition no chain can contain that minor premiss of an application of $\rightarrow E$ that is the conclusion of the minor sub-proof.) Furthermore, α cannot be the premiss of an introduction, for that would make α terminal within the maximal chain, contrary to the requirement that the terminal occurrence of A within a maximal A-chain be the major premiss of an elimination. So α must be a premiss (major or minor) of an elimination. But then the proof Π would contain a step of $\vee$-elimination immediately followed by another elimination, i.e., Π would not be shuffled, contrary to hypothesis.

Definition: A *normal* proof is a proof with no maximal chains.

This is the faithful generalization, to the setting of parallelized proofs, of the 'Prawitzian' notion of normal form (in the setting of serial proofs), that we mentioned earlier.

Definition: The *degree of a sentence* is the number of occurrences of logical operators within it.

Definition: The *degree of a chain* is the degree of the sentence it involves.

Definition: The *degree of abnormality of a proof* is determined as follows. Let n_d be the number of maximal chains of degree d within the proof. Let k be the highest of these degrees. (k must exist, since the proof is finite.) Then the degree of abnormality of the proof is the ordered k-tuple $< n_k,\ n_{k-1},\ \ldots,\ n_1 >$.

Degrees of abnormality are well-ordered as though one were 'counting to base ω'.

Definition: *The most egregious maximal chain* of a proof Π (call it γ_Π) is found as follows. Consider the maximal chains in Π. Certain ones among these will be of the maximal degree. Of the latter, consider those highest within Π, as judged by their lowest sentence-occurrences; and of those highest ones, take the leftmost. This is γ_Π.

Once a proof has been shuffled, with all its $\vee$-eliminations driven down as far as possible, one can set about applying reduction procedures—to be defined below—and *once again shuffling the results*, to convert the proof into normal form.

Definition: We shall designate by $\rho(\Pi)$ *the result of applying a single reduction procedure to the most egregious maximal chain* γ_Π within a *shuffled* proof Π. An application of a reduction procedure will be called a reduction. (The reduction procedures are listed below.)

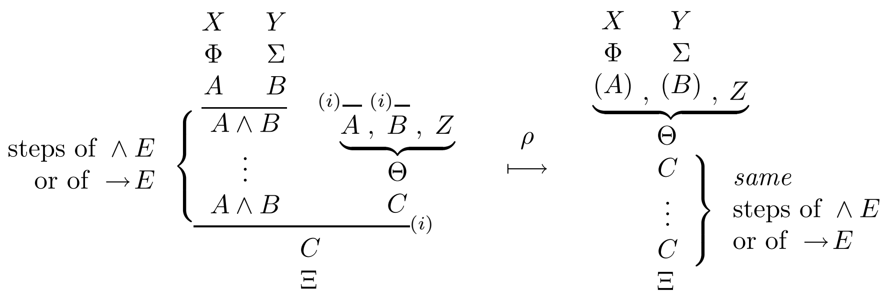
Note, then, that the operation ρ is defined *only* on shuffled proofs, and *only* with respect to the most egregious maximal chain within the proof concerned.

The reader is advised at this stage to look ahead to the pre-images and images of the transformation ρ listed below. The following brief clarificatory remarks are to be interpreted with an eye to those schemata.

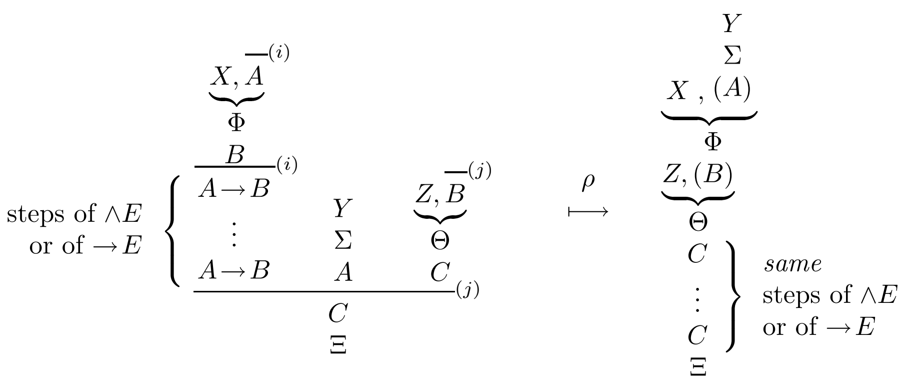
What is shown below is what happens, under the action of ρ, to the sub-tree of Π determined by the conclusion of the elimination that terminates (i.e. stands bottommost within) the chain γ_Π. If that sub-tree were to be excised from Π, then the residue would be what is called Ξ within each pre-image for ρ. The reduction-cases are determined by the logical operator dominant in the link-sentence A that occurs in the branch γ_Π; and by whether the topmost occurrence of A in the chain γ_Π is the conclusion of an introduction, or the conclusion of an application of the absurdity rule. In each case, the long curly parenthesis in the pre-image for the transformation ρ on Π indicates the extent of the chain γ_Π, which runs through the fragment of Π that is indicated by the three vertical dots. (We shall say more about this fragment presently.) The corresponding long curly parenthesis in the image $\rho(\Pi)$ indicates the extent of the isomorphic branch-segment that results from γ_Π by replacing each occurrence of the link-sentence within γ_Π by an occurrence of the simpler sentence shown on the right within the scope of the curly parenthesis. Thus the transform $\rho(\gamma_\Pi)$ has a different link-sentence than γ_Π. Remember that each occurrence of the link-sentence within the chain γ_Π stands as the conclusion of either $\wedge$-elimination or $\rightarrow$-elimination. (It cannot stand as the conclusion of $\vee$-elimination, because all $\vee$-eliminations have, *ex hypothesi*, been shuffled down below γ_Π in the pre-image for the transformation ρ.) All parts of the proof Π standing above any of the *other* premisses for these $\wedge$- and $\rightarrow$-eliminations—that is, premisses other than the various occurrences of the link-sentence involved in the chain γ_Π—remain undisturbed by the substitution effected, in the image, for the link-sentence whose occurrences make up γ_Π. That is why we can speak loosely of the '*same*' steps of $\wedge$- and of $\rightarrow$-elimination occurring in the image as occur, before that substitution, in the pre-image. The geometry of the branch-segment remains unchanged. Moreover, the reduction ρ in question leaves undisturbed the surrounding proof-fragment 'sideways connected with' the chain γ_Π. This proof-fragment, whose presence is signalled in each pre-image proof-schema by the vertical dots, consists of those parts of Π that stand above co-premisses of the $\wedge$- and/or $\rightarrow$-eliminations that propagate the chain γ_Π.

REDUCTION PROCEDURES
for maximal chains beginning with introductions

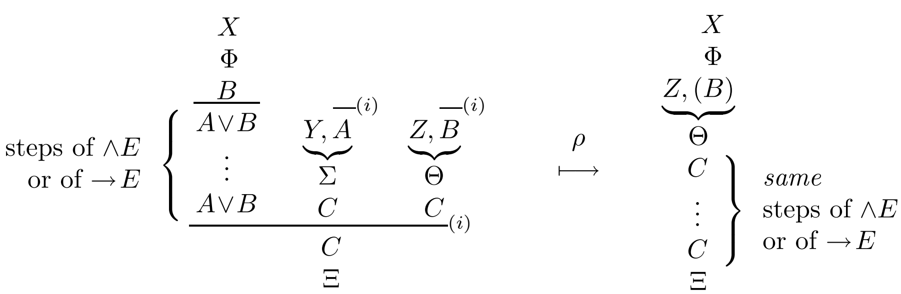
$$
\begin{array}{c}
\underbrace{X, \overline{A}^{(i)}} \\
\Phi \\
\text{steps of } \wedge E \text{ or of } \rightarrow E \left\{ \begin{array}{l} \dfrac{\bot}{\neg A}(i) \\ \vdots \\ \neg A \end{array} \right. \\
\end{array}
\;
\begin{array}{c} Y \\ \Sigma \\ A \end{array}
\quad \overline{\bot} \quad \Xi
\qquad \overset{\rho}{\longmapsto} \qquad
\begin{array}{c}
Y \\ \Sigma \\
\underbrace{X \;,\; (A)} \\
\Phi \\
\left. \begin{array}{c} \bot \\ \vdots \\ \bot \end{array} \right\} \begin{array}{l} \textit{same} \\ \text{steps of } \wedge E \\ \text{or of } \rightarrow E \end{array} \\
\Xi
\end{array}
$$



If the conclusion C of Θ is part of a maximal chain in the transform, then it would have to have been part of a maximal chain in the pre-image. In that case the degree of C cannot exceed that of $A \wedge B$. It follows that the reduction lowers the overall degree of abnormality of the proof.



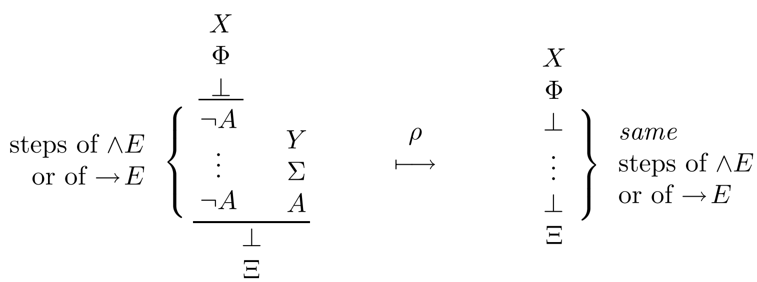
If the conclusion C of Θ is part of a maximal chain in the transform, then it would have to have been part of a maximal chain in the pre-image. In that case the degree of C cannot exceed that of $A \rightarrow B$. It follows that the reduction lowers the overall degree of abnormality of the proof.

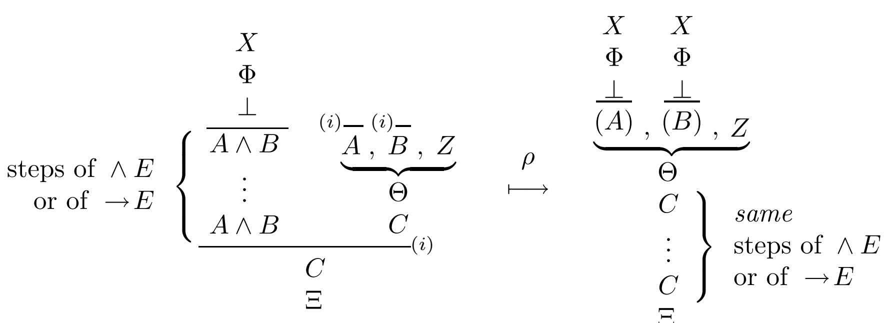


If the conclusion C of Θ is part of a maximal chain in the transform, then it would have to have been part of a maximal chain in the pre-image. In that case the degree of C cannot exceed that of $A \vee B$. It follows that the reduction lowers the overall degree of abnormality of the proof.

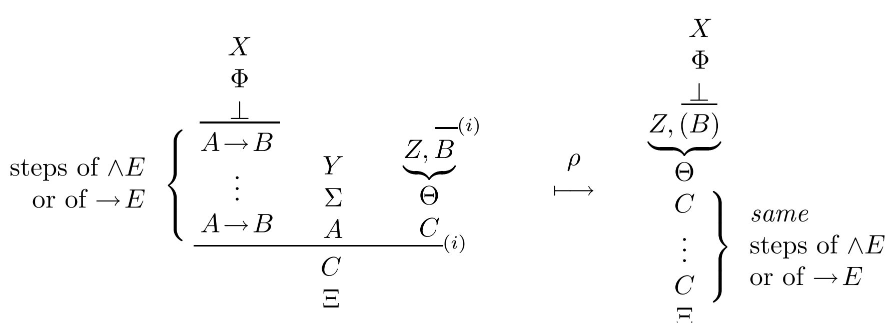
(The reduction shown above applies when the premiss for the step of $\vee I$ yielding $A \vee B$ is B. A similar reduction obtains when the premiss for that step is A.)

REDUCTION PROCEDURES
for maximal chains beginning with applications of the absurdity rule

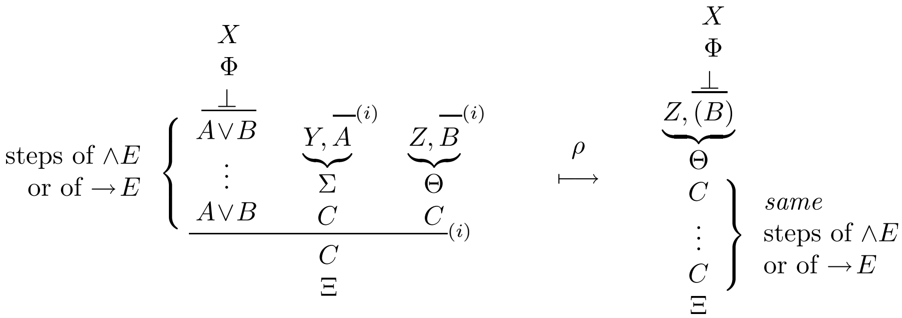




If the conclusion C of Θ is part of a maximal chain in the transform, then it would have to have been part of a maximal chain in the pre-image. In that case the degree of C cannot exceed that of $A \wedge B$. It follows that the reduction lowers the overall degree of abnormality of the proof.



If the conclusion C of Θ is part of a maximal chain in the transform, then it would have to have been part of a maximal chain in the pre-image. In that case the degree of C cannot exceed that of $A \rightarrow B$. It follows that the reduction lowers the overall degree of abnormality of the proof.



If the conclusion C of Θ is part of a maximal chain in the transform, then it would have to have been part of a maximal chain in the pre-image. In that case the degree of C cannot exceed that of $A \vee B$. It follows that the reduction lowers the overall degree of abnormality of the proof.

(The reduction is somewhat arbitrary in that it chooses Θ to place below Φ, when Σ could have been chosen instead.)

General observation about reductions: It is clear from how we chose γ_Π that none of the possibly replicated sub-proofs within $\rho(\Pi)$ will contain a maximal chain of equal or greater degree. Hence the reduction decreases *either* the number of maximal chains of highest degree within Π, *or* (in the case where γ was the sole maximal chain of highest degree) the highest degree of maximal chains. That is, the degree of abnormality of $\rho(\Pi)$ is lower than that of Π.

Normalization Theorem for Intuitionistic Logic. Given any shuffled intuitionistic proof Π of A from X, finitely many reductions followed by shufflings will turn Π into a normal intuitionistic proof of A from (some subset of) X.
Proof. In order to normalize Π proceed as follows. First, find γ_Π, the most egregious maximal chain in Π. Perform the appropriate reduction on it to obtain $\rho(\Pi)$. Then shuffle, if necessary, to obtain $\sigma(\rho(\Pi))$.

The shufflings that may be involved in the operation σ applied to $\rho(\Pi)$ will leave the degree of abnormality of $\sigma(\rho(\Pi))$ less than that of Π. For, any new maximal chain created by the further shuffling of $\rho(\Pi)$ will *either* have as its sentence one of the subsentences of the sentence in γ_Π, and hence be of lower degree than γ_Π; *or* have as its sentence one occurring in a maximal chain within a sub-proof for the introduction initiating γ_Π, and hence, because of the way γ_Π is chosen, be of lower degree than γ_Π.

It follows that finitely many reductions, followed if need be by shufflings, will eventually yield a proof that contains no maximal chains at all.

It is also clear, from inspection of the reduction procedures, that the undischarged assumptions of the resulting proof in normal form will be among those of Π. *QED*

6.3 Converting normal-form proofs into exposed normal form

In a proof in ordinary normal form, no major premiss for an elimination stands as the conclusion of any application of an introduction rule. In a proof in *exposed* normal form, the requirement is made even more exigent:

Definition. A proof is in *exposed normal form* just in case it has no major premiss for an elimination standing as the conclusion of *any* rule-application.

That is to say, in a proof in exposed normal form major premisses for eliminations *stand proud* at the tips of branches within the proof-tree. In general, some of them may be discharged by steps of inference 'lower down' within the proof; but none of them will be inferred *from* anything standing above it.

In order to help the reader appreciate the fact that there are proofs in normal form that are not in exposed normal form, here is an example:

$$\dfrac{\dfrac{A\to B \qquad A \qquad \dfrac{B\to(C\wedge D) \qquad \overline{B}^{(3)} \qquad \overline{C\wedge D}^{(1)}}{C\wedge D}{\scriptstyle (1)}}{C\wedge D}{\scriptstyle (3)} \qquad \dfrac{C\to E \qquad \overline{C}^{(2)} \qquad \overline{E}^{(4)}}{E}{\scriptstyle (4)}}{E}{\scriptstyle (2)}$$

Note that while there is no *maximal* chain in this proof (simply because it contains no introductions, nor any application of the absurdity rule), it does nevertheless contain a chain of length 3, whose link-sentence is $C\wedge D$.

Definition. For a proof Π in normal form ending with an elimination, the *spine* of Π is the unique longest segment of major premisses for eliminations terminating on the major premiss of the terminal elimination.

The notion of spine is from [14]. Note the insistence that the proof be in normal form, and end with an elimination; otherwise, the notion of spine is not uniquely defined. If the proof is in normal form, but does not end with an elimination, then it will have a (possibly empty) set of spines.

A spine must not be confused with a chain. For a chain always consists of occurrences of some same sentence (which we called its link-sentence); whereas spines will in general be made up of occurrences of different sentences. Thus, for example, in the normal-form proof above, the spine consists of the MPE $A\to B$ of the step of $\to$-elimination labelled (3), followed by the MPE $C\wedge D$ of the step of $\wedge$-elimination labelled (2).

The lengths of its (possibly various) maximal chains afforded a well-founded measure of the 'degree of abnormality' of a proof, before being normalized. Once it has been normalized, the lengths of its various spines will then afford a well-founded measure of the extent to which a (shuffled) normal-form proof still departs from what we shall call *exposed normal form*.

We now show how to turn any shuffled intuitionistic proof in normal form into one in *exposed* normal form. *The aim is to ensure that, after the exposure operation has been* *applied, every spine in the resulting proof consists of only one sentence occurrence.* Thus in an exposed normal form, every MPE will ‘stand proud’, with no proof-work above it—not even any applications of elimination rules. By way of illustration, here is the exposed normal form of the proof given above:

$$\dfrac{A\to B \quad A \quad \dfrac{B\to(C\wedge D)\quad \dfrac{}{B}{\scriptstyle(4)}\quad \dfrac{\dfrac{}{C\wedge D}{\scriptstyle(3)}\quad \dfrac{C\to E\quad \dfrac{}{C}{\scriptstyle(2)}\quad \dfrac{}{E}{\scriptstyle(1)}}{E}{\scriptstyle(1)}}{E}{\scriptstyle(2)}}{E}{\scriptstyle(3)}}{E}{\scriptstyle(4)}$$

The exposure operation, which we shall call ϵ, is defined inductively as follows. Remember that this operation applies only to proofs already in shuffled normal form.

Inductive Definition of the operation ϵ

For the simplest possible kind of proof, consisting of a single occurrence of the sentence A:

$$(A)^\epsilon =_{df} A$$

For a proof ending with an application of the absurdity rule:

$$\left(\dfrac{\Pi}{A}\right)\epsilon \quad =_{df} \quad \dfrac{\Pi^\epsilon}{A}$$

For a proof ending with an introduction:

$$\left(\dfrac{\Pi}{A}\,I\right)\epsilon \quad =_{df} \quad \dfrac{\Pi^\epsilon}{A}\,I$$

So far so simple. But now we have to consider the cases where the terminal step in Π is an elimination. Find the topmost sentence occurrence of the spine in Π. Possible cases for the pre-image are determined by the dominant operator of the sentence involved:

$$\dfrac{\neg A \qquad \begin{array}{c}\Sigma\\ A\end{array}}{\bot}$$

(Remember that $\bot$ cannot be a major premiss for an elimination!)

$$\dfrac{A\vee B \qquad \begin{array}{c}\overline{A}^{(i)}\\ \Sigma\\ C\end{array} \qquad \begin{array}{c}\overline{B}^{(i)}\\ \Phi\\ C\end{array}}{C}{\scriptstyle(i)}$$

(Remember that the proof is shuffled!)

$$\begin{array}{l}\dfrac{A \wedge B \qquad \begin{matrix} {}^{(i)}\overline{A}\;,\;\overline{B}^{(i)} \\ \Sigma \\ C \end{matrix}}{}\,(i) \\ \text{all E's}\left\{\begin{matrix} C \\ \vdots \\ \overline{D}^{\,E} \end{matrix}\right. \end{array} \qquad\qquad \begin{array}{l}\dfrac{A \to B \quad \begin{matrix} \Sigma \\ A \end{matrix} \quad \begin{matrix} \overline{B}^{(i)} \\ \Phi \\ C \end{matrix}}{}\,(i) \\ \text{all E's}\left\{\begin{matrix} C \\ \vdots \\ \overline{D}^{\,E} \end{matrix}\right. \end{array}$$

Having identified the possible forms that might be inputs for the method of exposure, we now complete its specification. The ϵ-transforms in the first two cases are easy:

$$\dfrac{\neg A \qquad \begin{matrix} \Sigma^{\epsilon} \\ A \end{matrix}}{\bot} \qquad\qquad \dfrac{A \vee B \quad \begin{matrix} \overline{A}^{(i)} \\ \Sigma^{\epsilon} \\ C \end{matrix} \quad \begin{matrix} \overline{B}^{(i)} \\ \Phi^{\epsilon} \\ C \end{matrix}}{C}\,(i)$$

The ϵ-transforms in the other two cases are as follows:

$$\dfrac{A \wedge B \qquad \begin{pmatrix} {}^{(i)}\overline{A}\;,\;\overline{B}^{(i)} \\ \Sigma \\ \{C\} \\ \vdots \\ \overline{D}^{\,E} \end{pmatrix}^{\epsilon}}{D}\,(i) \qquad\qquad \dfrac{A \to B \quad \begin{matrix} \Sigma^{\epsilon} \\ A \end{matrix} \quad \begin{pmatrix} \overline{B}^{(i)} \\ \Phi \\ \{C\} \\ \vdots \\ \overline{D}^{\,E} \end{pmatrix}^{\epsilon}}{D}\,(i)$$

That ends the inductive definition of the operation ϵ.

Note that ϵ *only re-orders* steps. The new sub-proofs to be ϵ'd (recursively, as indicated) are composed of existing steps, and involve no multiple copies. Hence the inductively defined exposure operation ϵ terminates.

Clearly, ϵ preserves both shuffled form and terminal character of a proof. It also preserves normality. But the result of ϵ is that no major premiss for an elimination is a conclusion of any step in the proof! Every MPE has been 'exposed', so as to stand proud, occupying a leaf node in the proof-tree. (Of course, such MPEs might themselves still be *discharged* by steps of inference lower down; but that does not alter the fact that they have been 'exposed' in the manner described.)

The foregoing discussion has established the following result.

Exposed Normal Form Theorem. Every intuitionistic proof can be transformed into a proof in exposed normal form of the same result.

7 Intuitionistic relevant logic

The formal display below is of a subtly, but importantly, different system of introduction and elimination rules for the intuitionistic logical connectives. This system

is what we have been calling *intuitionistic relevant* logic (IR). Explanations of our notational conventions will follow. Two important remarks must, however, be made at the outset.
(i) In all applications of the elimination rules stated below, the major premiss *stands proud*; that is, it is not the conclusion of any rule.
(ii) No applications of the absurdity rule are permitted.

	INTRODUCTION	ELIMINATION
$\neg$	$\dfrac{\begin{matrix}\Box\overline{}^{(i)} \\ A \\ \vdots \\ \bot\end{matrix}}{\neg A}\,(i)$	$\dfrac{\neg A \quad A}{\bot}$
$\wedge$	$\dfrac{A \quad B}{A \wedge B}$	$\dfrac{A \wedge B \quad \begin{matrix}{}^{(i)}\overline{}\Box\overline{}^{(i)} \\ \underbrace{A\ ,\ B} \\ \vdots \\ C\end{matrix}}{C}\,(i)$
$\vee$	$\dfrac{A}{A \vee B} \qquad \dfrac{B}{A \vee B}$	$\dfrac{A \vee B \quad \begin{matrix}\Box\overline{}^{(i)} \\ A \\ \vdots \\ \bot/C\end{matrix} \quad \begin{matrix}\Box\overline{}^{(i)} \\ B \\ \vdots \\ \bot/C\end{matrix}}{\bot/C}\,(i)$
$\rightarrow$	$\dfrac{\begin{matrix}\Diamond\overline{}^{(i)} \\ A \\ \vdots \\ B\end{matrix}}{A \rightarrow B}\,(i) \qquad \dfrac{\begin{matrix}\Box\overline{}^{(i)} \\ A \\ \vdots \\ \bot\end{matrix}}{A \rightarrow B}\,(i)$	$\dfrac{A \rightarrow B \quad A \quad \begin{matrix}\Box\overline{}^{(i)} \\ B \\ \vdots \\ C\end{matrix}}{C}\,(i)$

Some explanatory comments are now in order.

7.1 Liberalized proof by cases

The rule of proof by cases, or $(\vee E)$, looks unusual. By stating it graphically as we have, we are providing for the possibility that one of the case assumptions might lead to absurdity $(\perp)$. We are then permitted to bring down as the main conclusion whatever is concluded from the *other* case assumption. Thus the rule as stated is shorthand for the following possibilities regarding the sequence *Conclusion of first case-proof, Conclusion of second case-proof, ergo Main conclusion*:

C, C, *ergo* C (including, as a special case, $\perp$, $\perp$, *ergo* $\perp$);
C, $\perp$, *ergo* C; and
$\perp$, C, *ergo* C.

Liberalizing proof by cases in this way is entirely natural, given how we reason informally. Suppose one is told that $A \vee B$ holds, along with certain other assumptions X, and one is required to prove that C follows from the combined assumptions X, $A \vee B$. If one assumes A and discovers that it is inconsistent with X, one simply stops one's investigation of that case, and turns to the case B. If C follows in the latter case, one concludes C as required. One does *not* go back to the conclusion of absurdity in the first case, and artificially dress it up with an application of the absurdity rule so as to make it also 'yield' the conclusion C.

7.2 Vacuous v. non-vacuous discharge of assumptions

Note that in the rules for *IR* above we make some unusual demands on the *discharges of assumptions* allowed by certain of these rules. By contrast with minimal logic, we retain the permissive construal only for (the first half of) the rule of $(\rightarrow I)$. Neither minimal nor intuitionistic logic insists on non-vacuous discharge when applying the rule $(\neg I)$. The intuitionist logician has a reason of sorts for this omission: any application of $(\neg I)$ to infer $\neg A$ *without* having used A to obtain the preceding absurdity could simply be regarded instead as an application of the absurdity rule. The latter rule, however, is conspicuously absent from our relevant system *IR*, and for good reason. It leads to irrelevancies such as Lewis's first paradox $(A, \neg A : B)$. We are concerned to avoid this paradox in *IR*. Just as disagreeable is the negated-conclusion version of the paradox: $A, \neg A : \neg B$. Thus we have good reason to insist that $\neg I$ be applied only with non-vacuous discharge.

7.3 IR is contained in intuitionistic logic

No one who accepts intuitionistic logic can balk at any of the rules proposed above. Every one of them is either already primitive, or easily derivable, within the standard system of intuitionistic logic. (Note: in order to derive our newly liberalized forms of $(\rightarrow I)$ and $(\vee E)$, the intuitionistic logician will employ the absurdity rule!) It should be noted, however, that a *proof* in *IR* need not, in general, be a proof in intuitionistic logic. This is precisely because certain steps that are primitive in *IR* need to be *derived* in intuitionistic logic by special recourse to the absurdity rule.

The new discharge conventions, which impose the requirement that certain assumptions should have been used within a sub-proof, cannot incur any objection on grounds of *unsoundness*; for, when the requirement is met, the application of the rule in question is licit in the standard system. The only objection that could be raised, therefore, would be that the new requirements lead to some sort of *incompleteness* in the resulting system. But this objection also will be shown to be groundless, by means of a metatheorem already discussed above, and whose proof is to follow. In brief: whatever one might reasonably require intuitionistic logic to do (as a canon of mathematical and scientific inference), the system of intuitionistic relevant logic can do.

7.4 An important property of the rules of IR

Definition: When talking of forms of labelled trees, the notation

$$\frac{\Sigma_1 \ldots\ \Sigma_n}{A/\bot}$$

will be used as a complex schematic variable ranging over labelled trees that are either of the form

$$\frac{\Sigma_1 \ldots\ \Sigma_n}{A}$$

or of the form

$$\frac{\Sigma_1 \ldots\ \Sigma_n}{\bot}\ .$$

Take any non-trivial *IR*-proof Π with conclusion A. Its final step will be an introduction or an elimination. Thus Π will have one of the following forms:

$$\frac{\Sigma_1 \quad [\Sigma_2]}{A/\bot}{\scriptstyle (I)} \qquad\qquad \frac{\text{MPE} \quad \Sigma_1 \quad [\Sigma_2]}{A/\bot}{\scriptstyle (E)}$$

Let Σ^* represent any *IR*-proof at least as strong as the *IR*-proof Σ. Then the rules of *IR* have the following important property.

Lemma on Strengthening of Sub-Proofs. If Π is an I-proof of the form

$$\frac{\Sigma_1 \quad [\Sigma_2]}{A/\bot}{\scriptstyle (I)}$$

and, for given *IR*-proof Σ_i^* at least as strong as Σ_i, neither tree of the form

$$\frac{\Sigma_1^* \quad [\Sigma_2^*]}{A/\bot}{\scriptstyle (I)}$$

is an *IR*-proof, then for some i (= 1 or 2) the strengthened sub-proof Σ_i^* is at least as strong as Π; and if Π is an I-proof of the form

$$\frac{\text{MPE} \quad \Sigma_1 \quad [\Sigma_2]}{A/\bot}{\scriptstyle (E)}$$

and, for given *IR*-proof Σ_i^* at least as strong as Σ_i, neither tree of the form

$$\frac{\text{MPE} \quad \Sigma_1^* \quad [\Sigma_2^*]}{A/\bot}{\scriptstyle (E)}$$

is an *IR*-proof, then for some i (= 1 or 2) the strengthened sub-proof Σ_i^* is at least as strong as Π.
Proof. By inspection of the rules of *IR*, with special attention paid to the liberalized forms of ($\vee E$) and ($\rightarrow I$), and the various obligatory discharge requirements.

8 The motivation for our treatment

We have two goals in mind. The first is to show that there is a natural way to relevantize intuitionistic logic, involving no loss of completeness for mathematical and scientific purposes. Since this has largely been accomplished already in other publications, however, it is very much a secondary goal in this exposition. (See [19], *Autologic*, and [23], esp. ch. 10.)

Our primary goal in this paper is to demonstrate that the recommended approach to natural deduction obviates the need to resort to Gentzen's sequent calculi in order to prove results about the system.

Gentzen intended his system of natural deduction to '[come] as close as possible to actual reasoning' (loc. cit., p. 68). But, in order to prove his *Hauptsatz* about the existence of normal forms of proofs, he said (loc. cit., p. 69.)

> I had to provide a logical calculus especially suited to the purpose. *For this the natural calculus proved unsuitable.* [My emphasis.] For, although it already contains the properties essential to the validity of the *Hauptsatz*, it does so only with respect to its intuitionist form . . . .

It is important to realize that when Gentzen spoke of his *Hauptsatz* he had in mind a general kind of result, rather than a specific theorem concerning a particular formalism such as the sequent calculus. He characterized the *Hauptsatz* quite generally as saying that

> every purely logical proof can be reduced to a definite, though not unique, normal form. Perhaps we may express the essential properties of such a normal proof by saying: it is not roundabout. No concepts enter into the proof other than those contained in its final result, and their use was therefore essential to the achievement of that result.

One is invited to draw the conclusion that, had he been concerned only with intuitionistic logic, Gentzen might well have proved the *Hauptsatz* directly for his intuitionistic natural deduction system *NJ*, without inventing the corresponding sequent system *LJ*. It was only with the later work of Prawitz ([6], p. 50, Theorem I) however, that the normalizability of intuitionistic (and minimal) natural deductions was proved directly, without recourse to the sequent calculus. And indeed, in that work Prawitz had to make do, in the classical case, with a normalization theorem only for the system without the irksome operators $\vee$ and $\exists$. The normalization theorem for classical logic with all the usual operators primitive was obtained only much later, by Stålmarck in 1991.

When Gentzen first formulated the system of natural deduction ('*NJ*' for intuitionistic logic, '*NK*' for classical logic) he gave three examples (loc. cit., pp. 74–5) of pieces of informal reasoning that could be straightforwardly formalized as natural deductions. It is a great pity that his examples were so simple and so limited in number. A greater variety of even slightly more complex examples would have invited the thought that the serial forms of $(\wedge E)$ and of $(\rightarrow E)$ carried the cost, remarked on above, of replication of proof of their major premisses. Indeed, none of Gentzen's examples involved $(\rightarrow E)$. Moreover, the only example that involved $(\wedge E)$ had it applied to an assumption. Thus this major premiss for $(\wedge E)$ did not stand as the conclusion of a non-trivial sub-proof; so its repetition for the second application of $(\wedge E)$ (to obtain the other conjunct) would tend to go unremarked.

8.1 The method of extraction

The ban, in (*IR*), on various kinds of vacuous discharge that are permitted in intuitionistic natural deduction raises the question as to what exactly the relationship is between intuitionistic natural deduction (*NJ*) and the system (*IR*) we have presented above. (For convenience, we take both (*NJ*) and (*IR*) to agree at least in having all elimination rules parallelized; the question is only what difference obtains as a result of (a) the differing attitudes to vacuous discharge, and (b) the use, in the system (*NJ*), of the absurdity rule.) The relationship is very straightforward.

Theorem. Every intuitionistic proof in exposed normal form can be converted into an *IR*-proof that is at least as strong.

Proof. Suppose Π is an intuitionistic proof in exposed normal form. The *IR*-proof Π^* can be *extracted* from Π by induction on the construction of Π. For the basis, $A^* =_{df} A$. If Π ends with an application of the absurdity rule, and so is of the form

$$\dfrac{\begin{matrix}\Sigma\\ \bot\end{matrix}}{A}$$

then $\Pi^* =_{df} \Sigma^*$. If Π is of the form

$$\dfrac{\Sigma_1 \quad [\Sigma_2]}{A}{\scriptstyle (I)}$$

then Π^* will be

$$\dfrac{\Sigma_1^* \quad [\Sigma_2^*]}{A}{\scriptstyle (I)}$$

if this is an *IR*-proof; otherwise, by the Lemma on Strengthening Sub-Proofs, for some i (= 1 or 2), Π^* can be set equal to Σ_i^*, which is at least as strong as Π. Finally, if Π is of the form

$$\dfrac{\text{MPE} \quad \Sigma_1 \quad [\Sigma_2]}{A/\bot}{\scriptstyle (E)}$$

then Π^* will be

$$\dfrac{\text{MPE} \quad \Sigma_1^* \quad [\Sigma_2^*]}{A/\bot}(E)$$

if this is an *IR*-proof; otherwise, by the Lemma on Strengthening Sub-Proofs, for some i (= 1 or 2), Π^* can be set equal to Σ_i^*, which is at least as strong as Π. Since Π is in exposed normal form, the possibilities we have considered are exhaustive. *QED*

It is clear that the extraction of Π^* from Π can be performed in linear time. There will in general be more than one *IR*-proof that can be extracted from Π, and that establishes a result at least as strong as that of Π. For note that there can be cases where *both* Σ_1^* *and* Σ_2^* are eligible to be set equal to Π^*. The extraction procedure is in general non-deterministic.

The extracted version of the proof of disjunctive syllogism given above would be

$$\dfrac{A \vee B \qquad \dfrac{\neg A \qquad \overline{A}^{(1)}}{\bot} \qquad \overline{B}^{(1)}}{B}(1)$$

8.2 Ultimate normal forms

Putting together our earlier results, we now have our central result.

Main Theorem. Any intuitionistic proof can be converted into an *IR*-proof that is at least as strong.

Proof. Let Π be an intuitionistic proof, with conclusion A and set X of undischarged assumptions. The master sequence of operations that we have provided is as follows:

$$\begin{matrix} X \\ \Pi \\ A \end{matrix} \overset{\text{shuffle}}{\longmapsto} \begin{matrix} X \\ \Pi^{\sigma} \\ A \end{matrix} \overset{\text{reduce/shuffle}}{\longmapsto} \begin{matrix} X' \\ \Pi^{\sigma\rho} \\ A \end{matrix} \quad (\text{where } X' \subseteq X)$$

$$\overset{\text{expose}}{\longmapsto} \begin{matrix} X' \\ \Pi^{\sigma\rho\epsilon} \\ A \end{matrix} \overset{\text{extract}}{\longmapsto} \begin{matrix} X'' \\ \Pi^{\sigma\rho\epsilon *} \\ A/\bot \end{matrix} \quad (\text{where } X'' \subseteq X')$$

Definition. We shall call $\Pi^{\sigma\rho\epsilon *}$ the *ultimate normal form* of the intuitionistic proof Π.

The proof $\Pi^{\sigma\rho\epsilon *}$ will be a proof constructed according to the rules of *IR*. In some cases it will be correct according to the rules of I; in some cases it will not (see the following Observation). But in all cases its overall result will be intuitionistically provable. Such a normal form $\Pi^{\sigma\rho\epsilon *}$ satisfies the following conditions:

- every major premiss for an elimination stands proud (which entails, among other things, that $\vee$-eliminations are indeed shuffled down as far as possible);
- there are no applications of the absurdity rule;
- every application of any one of $(\neg I)$, $(\wedge E)$, $(\vee E)$, $(\to I)$ (second half) or $(\to E)$ involves *non-vacuous* discharge of assumptions; and

- the result established is at least as strong as that established by Π.

Observation. Ironically, the ultimate normal form of an intuitionistic proof might not itself be an intuitionistic proof. Example:

$$\dfrac{\dfrac{\neg A \quad \overline{A}^{(1)}}{\bot}}{A \rightarrow B}{(1)}$$

But, as already remarked, every proof in ultimate normal form is a proof in the system of intuitionistic relevant logic.

9 Sequent rules for IR

Besides the natural deduction system above for *IR*, there is also the following sequent system.

The *rule of initial sequents* is that any sequent of the form $A{:}A$ is a sequent proof. More complex sequent proofs are then built up inductively by means of the following *logical rules* for introducing logical operators on the right, and on the left, of conclusion-sequents. *Note that no provision is made for the usual structural rules of cut and weakening. In the sequent system for IR, the rule of initial sequents is the* **only** *structural rule.*

THE SEQUENT RULES FOR *IR*

RIGHT	LEFT
$\dfrac{X, A :}{X : \neg A}$ A not in X	$\dfrac{X : A}{X, \neg A :}$
$\dfrac{X : A \quad Y : B}{X, Y : A \wedge B}$	$\dfrac{X : C}{X \setminus \{A, B\}, A \wedge B : C}$ $X \cap \{A, B\}$ non-empty
$\dfrac{X : A}{X : A \vee B} \qquad \dfrac{X : B}{X : A \vee B}$	$\dfrac{X, A : Z \quad Y, B : W}{X, Y, A \vee B : Z, W}$ A not in $X \cup Y$, B not in $X \cup Y$; $Z \cup W$ at most a singleton
$\dfrac{X, [A] : [B]}{X : A \rightarrow B}$ A not in X; and the brackets mean that either A or B occurs	$\dfrac{X : A \quad Y, B : C}{X, Y, A \rightarrow B : C}$ B not in Y

in the premiss sequent, in
the way indicated

10 Natural deductions in *IR* ‘are’ cut-free, weakening-free sequent proofs

Theorem. Every natural deduction in *IR* ‘is’ (modulo its twigs for MPEs, and re-labeling of nodes with sequents rather than sentences) a sequent proof in *IR* of the same result.
Proof. Suppose Θ is a natural deduction in *IR* that establishes the conclusion A from the set X of undischarged assumptions. Consider Θ as a tree whose nodes are labelled by sentences. Focus on the sub-tree $\Theta^{\#}$ that consists of all nodes of Θ except for those labelled by major premisses for eliminations. Re-label each node in $\Theta^{\#}$ by the deducibility statement (sequent) established by that point within Θ. The result will be a *sequent proof* in *IR* of the sequent $X : A$. *QED*

In other words, a sequent proof in *IR can be read directly off* a natural deduction in *IR* simply by pruning away the twigs whose leaves are major premisses for eliminations, and by being assiduous in one's deductive bookkeeping at all other points within the proof-tree.

Conversely, we have the following result.

Theorem. Any cut-free, weakening-free sequent proof Θ of intuitionistic logic ‘is’ (via the addition of twigs for MPEs and the re-labeling of nodes with sentences rather than sequents) a natural deduction $\Theta^{\dagger}$ in *IR*.
Proof. First, note that Θ, being cut-free and weakening-free, will count as a sequent proof of intuitionistic relevant logic (*IR*), as given by the preceding sequent rules.

The conversion of Θ to $\Theta^{\dagger}$ is effected as follows. Initial sequents (of the form $A : A$) within Θ are replaced by occurrences of A. Then, below those leaf nodes of the proof-tree, sequent steps that introduce logical operators on the *right* are rendered as *introductions*, while sequent steps that introduce logical operators on the *left* are rendered as *eliminations*, with their major premisses supplied on an extra twig. When a conclusion-sequent lacks, in its antecedent, some sentence in the antecedent of a premiss-sequent, treat the step as *discharging* the sentence in question at its occurrences as an assumption above. *QED*

In other words, a natural deduction in *IR can be read directly off* a cut-free, weakening-free intuitionistic sequent proof simply by grafting on twigs whose leaves are major premisses for eliminations, and by re-labelling each node with the conclusion that has been established by that point. (The assumptions on which it depends will be the sentences labelling the nodes that are undischarged by that stage.)

Our investigations have therefore revealed that the idea of a (cut-free, weakening-free) sequent proof is already implicit in a conception of a certain kind of normal form for natural deductions. In the case of intuitionistic logic at least, Gentzen need not have resorted to the sequent calculus as an *alternative* to the system of natural deduction, in order to prove his *Hauptsatz*. It can be proved directly for parallelized natural deduction, and with the added benefit of securing deductive relevance along the way.

Indeed, the proof of the *Hauptsatz* for the sequent system can now proceed by appeal to the (ultimate) normalization theorem for the system of natural deduction. First one would establish the following two lemmata.

Lemma (α). Any cut-free intuitionistic sequent proof can be turned into a cut-free, weakening-free sequent proof (that is, a sequent proof in *IR*) of a result at least as strong.
Proof. The inductive proof of (α) is straightforward, given that the sequent rules for the logical operators in *IR* are framed in the foregoing liberal way, allowing for the merging of non-identical antecedents and succedents. This liberality allows one to cope with such strengthenings as might be effected on the results of subordinate proofs for the application of a logical rule. *QED*

Lemma (β). Any weakening-free intuitionistic sequent proof Π whose only cut is terminal can be turned into a cut-free weakening-free sequent proof of a result at least as strong.
Proof. Suppose the intuitionistic sequent proof Π in question is

$$\dfrac{\begin{array}{cc}\Pi_1 & \Pi_2\\ X:A & A,Y:B\end{array}}{X,Y:B}$$

Note that both Π_1 and Π_2 are cut-free and weakening-free, and hence are proofs in *IR*. Turn them into natural deductions $\Pi_1^\dagger$ and $\Pi_2^\dagger$ in ultimate normal form. Graft these natural deductions together:

$$\begin{array}{l}X\\ \Pi_1^\dagger\\ (A),Y\\ \quad\Pi_2^\dagger\\ \quad B\end{array}$$

The resulting array, of course, need not (in general) be in normal form. Nor, strictly speaking, will it necessarily be a properly formed natural deduction, in either of the senses here employed. For, first, if $\Pi_1^\dagger$ is non-trivial, and any of the indicated assumption-occurrences of A in $\Pi_2^\dagger$ is a major premiss for an elimination, then the array will not count as a natural deduction in *IR*. Secondly, the daggered sub-proofs might contain applications of rules of *IR* (such as liberalized ($\vee I$), or the first half of ($\rightarrow I$)) that do not count as applications of any primitive rule of intuitionistic logic.

We remedy this shortcoming by modifying the array so that it will count as an intuitionistic natural deduction. To this end, interpolate applications of the absurdity rule wherever necessary in order to qualify this array technically as a natural deduction (albeit possibly abnormal) in intuitionistic logic. Then, by the foregoing ultimate normalization theorem, turn the result into ultimate normal form—call it Σ. This ultimate normal form Σ will of course be a natural deduction in *IR*, and will establish a result at least as strong as the original. Now construct the cut-free, weakening-free sequent proof $\Sigma^{\#}$ according to the method described above. *QED*

Putting together weakening-elimination from cut-free sequent proofs (α) and cut-elimination from weakening-free sequent proofs (β), one obtains the following.

Theorem (Hauptsatz for intuitionistic sequent calculus). Every sequent proof in intuitionistic logic can be turned into a cut-free, weakening-free sequent proof (that is, a proof in IR) of a result at least as strong as the original.
Proof. Work down the given proof-tree from its leaf-nodes, applying (α) or (β) as called for, whenever one encounters an application of either weakening or cut, respectively.

11 Conclusion

We have focused here on propositional logic, but the methods generalize smoothly to first-order logic with the universal and existential quantifiers. The foregoing considerations reveal that natural deductions, suitably normalized, are thoroughly homologous to sequent proofs built up by means only of the logical rules. The homology is perfect in the case of intuitionistic relevant proofs. In the intuitionistic case, at least, the proof of Gentzen's *Hauptsatz* could have been conducted entirely within the framework of natural deduction. There is really no 'gap' needing to be bridged between the system of natural deduction and the sequent system.

We conclude this discussion by suggesting that Gentzen's 'oversights' could be said to have been twofold: (i) he did not provide for parallel forms of ($\wedge E$) and ($\rightarrow E$) in the system of natural deduction; and (ii) he did not place weakening-elimination on a par with cut-elimination. Had Gentzen done both these things, he could well have succeeded in combining a concern for transitivity of deduction with a concern for the *relevance* of premisses to conclusions deduced from them. Moreover, the difference between a natural deduction (in *IR*) and the corresponding sequent proof would have been reduced to their being slightly different labellings of what was essentially one and the same proof-tree (modulo the extra twigs for MPEs within a natural deduction in exposed normal form). In a sequent proof, each node is labelled with the whole sequent that has been established thus far. In a natural deduction, by contrast, each node is labelled (more economically) only by the *conclusion* that has been established thus far; and it is left to the reader effectively to read off from the proof-tree the undischarged assumptions on which that conclusion depends.

References

[1] Gerhard Gentzen. Untersuchungen über das logische schliessen. *Mathematische Zeitschrift*, I, II:176–210, 405–431, 1934, 1935. Translated as 'Investigations into Logical Deduction', in *The Collected Papers of Gerhard Gentzen*, edited by M. E. Szabo, North-Holland, Amsterdam, 1969, pp. 68–131.

[2] Felix Joachimski and Ralph Matthes. Short proofs of normalization for the simply-typed λ-calculus, permutative conversions and Gödel's $\top$. *Archiv für mathematische Logik*, 2002. forthcoming.

[3] E. G. K. López-Escobar. Standardizing the N systems of gentzen. In X. Caicedo and C. H. Montenegro, editors, *Models, Algebras and Proofs*, volume 203 of *Lecture Notes in Pure and*

Applied Mathematics, pages 411–434. Marcel Dekker Inc., New York, 1999.

[4] S. Negri and J von Plato. *Structural Proof Theory*. Cambridge University Press, 2001.

[5] L Paulson. *Isabelle: a generic theorem prover*. Springer-Verlag, Berlin, New York, 1994.

[6] Dag Prawitz. *Natural Deduction: A Proof-Theoretical Study*. Almqvist & Wiksell, Stockholm, 1965.

[7] Greg Restall. *An Introduction to Sub-Structural Logics*. Routledge, London, New York, 2000.

[8] Peter Schroeder-Heister. A natural extension of natural deduction. *Journal of Symbolic Logic*, 49:1284–1300, 1984.

[9] Neil Tennant. Entailment and proofs. *Proceedings of the Aristotelian Society*, LXXIX:167–189, 1979.

[10] Neil Tennant. A proof-theoretic approach to entailment. *Journal of Philosophical Logic*, 9:185–209, 1980.

[11] Neil Tennant. Perfect validity, entailment and paraconsistency. *Studia Logica*, XLIII:179–198, 1984.

[12] Neil Tennant. Minimal logic is adequate for popperian science. *British Journal for Philosophy of Science*, 36:325–329, 1985.

[13] Neil Tennant. Anti-realism and choice of logic. *Untersuchungen zur Logik und zur Methodologie*, 3:39–54, 1986.

[14] Neil Tennant. Natural deduction and sequent calculus for intuitionistic relevant logic. *Journal of Symbolic Logic*, 52:665–680, 1987.

[15] Neil Tennant. *Anti-Realism and Logic: Truth as Eternal*. Clarendon Library of Logic and Philosophy, Oxford University Press, 1987b.

[16] Neil Tennant. Truth table logic, with a survey of embeddability results. *Notre Dame Journal of Formal Logic*, 30:459–484, 1989.

[17] Neil Tennant. *Autologic*. Edinburgh University Press, 1992.

[18] Neil Tennant. Automated deduction and artificial intelligence. In R. Casati, B. Smith, and G. White, editors, *Philosophy and the Cognitive Sciences: Proceedings of the 16th International Wittgenstein Colloquium*, pages 273–286. Hölder-Pichler-Tempsky, Vienna, 1994.

[19] Neil Tennant. Intuitionistic mathematics does not need *ex falso quodlibet*. *Topoi*, pages 127–133, 1994.

[20] Neil Tennant. On maintaining concentration. *Analysis*, pages 143–152, 1994.

[21] Neil Tennant. Transmission of truth and transitivity of proof. In Dov Gabbay, editor, *What is a Logical System?*, pages 161–177. Oxford University Press, 1994.

[22] Neil Tennant. Delicate proof theory. In Jack Copeland, editor, *Logic and Reality: Essays on the Legacy of Arthur Prior*, pages 351–385. Oxford University Press, 1996.

[23] Neil Tennant. *The Taming of The True*. Oxford University Press, 1997.

[24] Neil Tennant. Negation, absurdity and contrariety. In Dov Gabbay and Heinrich Wansing, editors, *What is Negation*, pages 199–222. Kluwer, Dordrecht, 1999.

[25] Alasdair Urquhart. The undecidability of entailment and relevant implication. *Journal of Symbolic Logic*, 49:1059–1073, 1984.

[26] Jan von Plato. A problem of normal form in natural deduction. *Mathematical Logic Quarterly*, 46(1):121–4, 2000.

[27] Jan von Plato. Natural deduction with general elimination rules. *Archive for Mathematical Logic*, 40:541–567, 2001.

[28] Heinrich Wansing. Editorial introduction: Special issue on consequence relations and sequent systems. *Logic Journal of the IGPL*, 6(3):377–381, 1998.

Received 7 May, 2001.