Jan von Plato

FROM GENTZEN TO JASKOWSKI AND BACK: ALGORITHMIC TRANSLATION OF DERIVATIONS BETWEEN THE TWO MAIN SYSTEMS OF NATURAL DEDUCTION

Abstract

The way from linearly written derivations in natural deduction, introduced by Jaskowski and often used in textbooks, is a straightforward root-first translation. The other direction, instead, is tricky, because of the partially ordered assumption formulas in a tree that can get closed by the end of a derivation. An algorithm is defined that operates alternatively from the leaves and root of a derivation and solves the problem.

Keywords: proof systems, linear natural deduction, Gentzen, Jaśkowski

1. Linear proofs

Proofs are conventionally given as texts, i.e., linear successions of sentences. Each step of deduction can refer to one or several preceding sentences as a justification. With the advent of proof systems in the 1930s, a tree form was taken into use in which the premises of a deductive step stand on an inference line immediately above the conclusion and need not be singled out separately. In such proof trees, the deductive dependences of formulas are uniquely displayed through the top formulas of a subtree determined by a given formula in the tree. I shall define a translation algorithm for the conversion of tree proofs into linear proofs. This algorithm may look somewhat complicated in an abstract formulation, but when put to use, it is easily memorized and there is no difficulty in its application. The algorithm

is illustrated through *solutions of problems* rather than theorems, marked by the traditional *QEF, quod erat faciendum*.

I shall use a system of linear natural deduction with a handsome bracket notation for the handling of temporary hypotheses, as found in the Swedish compendium of Dag Prawitz, *ABC i Symbolisk Logik*. The system derives from previous similar systems of Irving Copi and Frederic Fitch that may go back to the work of Jaskowski (1934). The derivation below is self-explanatory of the rule system. Each line is either an assumption, a hypothesis to be closed, or a formula derived from previous ones, with the rule and line or lines of the premisses written next to the formula. The order in rule $\supset E$ is that the major premiss comes first.

Example derivation in linear natural deduction: $A \supset B \vdash \neg(A \& \neg B)$

$$
\begin{array}{lll}
1. & A \supset B & \text{assumption} \\
2. & A \& \neg B & \text{hypothesis: goal } \bot \\
3. & A & 2, \& E \\
4. & \neg B & 2, \& E \\
5. & B & 1, 3, \supset E \\
6. & \bot & 4, 5, \supset E \\
\hline
7. & \neg(A \& \neg B) & 2\text{--}6, \supset I \\
\end{array}
$$

The linear variety of natural deduction allows the construction of derivations in steps, one after the other, which feature makes it the accessible way to introduce the topic as exemplified by von Plato (2013). On the other hand, a formulation of rules for disjunction and existence elimination in a pure form requires modifications of the bracket notation that appear somewhat *ad hoc*. Moreover, there is the grave defect that what depends on what is not counted: Each formula apparently depends on the open assumptions above it. This situation is corrected in Gentzen's original system of natural deduction in which derivations are arranged in a tree form, with the following properties:

1. Each formula occurrence is either an initial formula (an assumption or temporary hypothesis) or the conclusion of exactly one logical rule.
2. Each formula occurrence is either the conclusion of the whole derivation or the premiss of exactly one logical rule.

3. Each formula occurrence in a derivation defines a subderivation the open assumptions of which are those initial formulas of the subtree that have not been closed by the formula occurrence.

The subderivations can be composed in new ways as in proofs of normalization that don't work in the linear variety; the latter does not have the properties that are needed for the composition of derivations. Tree derivations were in practice a novelty with Gentzen and their widespread use in logic derives from his doctoral thesis (1934–35). He took the idea over from the work of Paul Hertz of the 1920s. The way from linear to tree derivations is easily established:

Table 1. Translation of linear derivations to tree form
1. *Write down the endformula and a line above it.*
2. *Write next to the line the rule that was used in concluding the endformula. If it was $\supset I$, write after the rule the number of the line on which the hypothesis closed by the rule occurred.*
3. *Write above the line, from left to right, the formula or formulas that correspond to the numbers that justified the application of the rule. If it was $\supset I$, write the consequent of the implication.*
4. *Repeat the above until you come along each branch of the derivation tree to an assumption. If it is temporary, i.e., a hypothesis, write the number of its line above it.*

PROBLEM 1. *To apply the translation algorithm to the linear derivation that establishes $A \supset B \vdash \neg(A \,\&\, \neg B)$.*

Solution.

$$\cfrac{\cfrac{\overset{2}{A\,\&\,\neg B}}{\neg B}\,{}_{\&E_2} \quad \cfrac{A \supset B \quad \cfrac{\overset{2}{A\,\&\,\neg B}}{A}\,{}_{\&E_1}}{B}\,{}_{\supset E}}{\cfrac{\bot}{\neg(A\,\&\,\neg B)}\,{}_{\supset I,2}}$$

QEF.

The number 2 is found also above the hypothesis $A \,\&\, \neg B$, to indicate which hypothesis is closed where in a derivation. That assumption was used twice in the linear variant, and the formula appears correspondingly with two occurrences in the tree. The closed assumption is an initial formula. – The

only thing that reveals the origin of the derivation as a machine-produced result is the label 2 in place of 1 that would haven been chosen by a living logician who synthesizes such a derivation.

Note that there is no need to write, next to the rules, where their premisses come from, because we have by the translation:

> *In a derivation tree, the premisses of a rule stand directly above and the conclusion directly below the inference line of the rule.*

2. Linearization of tree derivations

The linearization of tree derivations is somewhat tricky because of a condition that has to be respected:

> **Bracket condition.** *In a linear derivation, the latest open assumption has to be closed first.*

The condition can be expressed in graphical terms as: The brackets must not cross each other. They can be nested or consecutive. We shall look at an example to see what is needed:

$\vdash A \& B \supset \neg\neg A \& \neg\neg B$

$$
\cfrac{
\cfrac{
\cfrac{\overset{1}{\neg A} \quad \cfrac{\overset{3}{A \& B}}{A} \,{}_{\& E_1}}{\bot} \,{}_{\supset E}
}{\neg\neg A} \,{}_{\supset I,1}
\qquad
\cfrac{
\cfrac{\overset{2}{\neg B} \quad \cfrac{\overset{3}{A \& B}}{B} \,{}_{\& E_2}}{\bot} \,{}_{\supset E}
}{\neg\neg B} \,{}_{\supset I,2}
}{
\cfrac{\neg\neg A \& \neg\neg B}{A \& B \supset \neg\neg A \& \neg\neg B} \,{}_{\supset I,3}
} \,{}_{\& I}
$$

In a linear derivation, the assumption $A \& B$ needs to be closed *after* the assumptions $\neg A$ and $\neg B$ have been closed. Thus, one has to look at the tree derivation at times down from the initial formulas at the leaves of the derivation, at other times up from the endformula.

Table 2. Linearization of tree derivations

1. *Number the open assumptions in a tree, from left to right in the obvious order, with each repetition of an assumption getting the same number, and let these numbers be 1 to n.*
2. *Continue the numbering of the formulas from formula 1 down until a two-premiss rule is reached, and repeat this in left to right order until formula n, with the last formula numbered being m.*

> 3. *If both branches of a two-premiss rule, always from left to right, have been numbered, continue from the conclusion until a two-premiss rule with an unnumbered branch is reached.*
> 4. *From the root of the derivation, find a downmost leftmost rule instance that closes a hypothesis, and number its occurrences from the previously reached number on, and number the formulas down from these as in 2. Topmost unnumbered formulas are now conclusions of two-premiss rules.*
> 5. *Continue numbering formulas in the branches, left to right, until you come to a rule instance with a branch that has an unnumbered hypothesis. Repeat 3 and 4.*

With the formulas numbered, it is clear how a linear derivation with the bracket notation is produced by writing the formulas in succession and adding the rules and justifications and brackets.

The full statement of the translation algorithm may look somewhat forbidding, but the matter is quite clear once it is seen in action:

PROBLEM 2. *To translate the tree derivation of $A \& B \supset \neg\neg A \& \neg\neg B$ in a linear form.*

Solution. The numbering produced for our example tree derivation is as follows, counting that there are no open assumptions, by which clause 4 is applied first, to be followed by clause 2:

$$
\cfrac{
\cfrac{
\overset{1}{\textbf{4. } \neg A} \quad \cfrac{\overset{3}{\textbf{1. } A \& B}}{\textbf{2. } A}{\scriptstyle \& E_1}
}{\cfrac{\textbf{5. } \bot}{\textbf{6. } \neg\neg A}{\scriptstyle \supset I,1}}{\scriptstyle \supset E}
\qquad
\cfrac{
\overset{2}{\textbf{7. } \neg B} \quad \cfrac{\overset{3}{\textbf{1. } A \& B}}{\textbf{3. } B}{\scriptstyle \& E_2}
}{\cfrac{\textbf{8. } \bot}{\textbf{9. } \neg\neg B}{\scriptstyle \supset I,2}}{\scriptstyle \supset E}
}{
\cfrac{\textbf{10. } \neg\neg A \& \neg\neg B}{\textbf{11. } A \& B \supset \neg\neg A \& \neg\neg B}{\scriptstyle \supset I,3}
}{\scriptstyle \& I}
$$

A linear derivation results by the writing of these formulas in the numbered succession, with the rule symbols, justification numbers, and brackets coming out in a unique way out of the numbered tree derivation:

$$
\begin{array}{lll}
\quad\;\; 1. & A \,\&\, B & \text{hypothesis: goal } \neg\neg A \,\&\, \neg\neg B \\
\quad\;\; 2. & A & 1, \&E \\
\quad\;\; 3. & B & 1, \&E \\
\quad\;\; 4. & \neg A & \text{hypothesis: goal } \bot \\
\quad\;\; 5. & \bot & 4, 2, \supset E \\
\quad\;\; 6. & \neg\neg A & 4\text{--}5, \supset I \\
\quad\;\; 7. & \neg B & \text{hypothesis: goal } \bot \\
\quad\;\; 8. & \bot & 7, 3, \supset E \\
\quad\;\; 9. & \neg\neg B & 7\text{--}8, \supset I \\
\quad\; 10. & \neg\neg A \,\&\, \neg\neg B & 6, 9, \&I \\
\quad\; 11. & A \,\&\, B \supset \neg\neg A \,\&\, \neg\neg B & 1\text{--}10, \supset I
\end{array}
$$

QEF.

I have found only one earlier attempt at finding the correct translation from tree derivations to linear derivations, that of Curry's 1965 "Remarks on inferential deduction." He did not pay attention to the intricacies in the order of closing assumptions, and his attempt can be failed by a counterexample.

Tree derivations, man-made or those produced through a system of proof search, often grow too broad to be printable or displayable on a screen. The translation algorithm here defined could perhaps be used for the linear display of such proofs. By the correspondence of natural deduction and sequent calculus, the linearization can be adapted even to the latter.

The translation can be extended to classical natural deduction with the standard rule of indirect proof. That gives a complete system, whereas we have to see how disjunction and existence can be handled in intuitionistic logic:

3. Adding disjunction and existence

Disjunction and existence elimination are somewhat awkward to formulate in the linear form of natural deduction. In Prawitz' manual, one derives $A \supset C$ and $B \supset C$ separately, then draws the conclusion C from the major premiss $A \vee B$, and analogously for existence: One proves $\forall y(A(y) \supset C)$ and then makes the conclusion C from the existential formula $\exists x A(x)$. In both cases, other logical operations than the one of the major premiss

come into play. With existence elimination, the usual informal practice in a linear proof text is, whenever an existential formula $\exists x A(x)$ is put into use, to write after the formula occurrence an instance $A(y)$ with an eigenvariable. This practice can be accommodated into the linear format: the subderivation from $A(y)$ to C will begin with a line under which the hypothesis $A(y)$ is written, as in the example:

1.	$\exists x A(x)$	assumption
2.	$A(y)$	hyp for $\exists E, 1$
3.	$A(y) \vee B(y)$	$2, \vee I$
4.	$\exists x(A(x) \vee (B(x)))$	$3, \exists I$
5.	$\exists x(A(x) \vee (B(x)))$	$1,\ 2\text{--}4,\ \exists E$

The hypothesis mentions the line of the major premiss. When the bracket is completed to close the hypothesis, the line of the major premiss is again indicated. The eigenvariable condition is immediate.

When a linear derivation with existence elimination is translated into a tree derivation, the inference line above the existence elimination will have first the major premiss, then the subderivation of the minor premiss from inside the bracket.

Disjunction presents a problem with its two minor premisses. A two-column subderivation notation as in the following example is handiest:

1. $A \vee B$	assumption			
2. A	hyp $\vee E, 1$		4. B	hyp $\vee E, 1$
3. $B \vee A$	$2, \vee I$		5. $B \vee A$	$4, \vee I$
6. $B \vee A$	$1,\ 2\text{--}3,\ 4\text{--}5,\ \vee E$			

With nested disjunction eliminations, there will be still a corresponding multiplication of doubling of columns so that derivations can grow too broad to be printable.

In the translation to tree form, the major premiss is written at left above the inference line of the disjunction elimination, then the derivation of the first minor premiss followed by the second.

4. Linearization of sequent derivations

Sequent calculi that support root-first proof systems are useful devices, but the resulting sequent derivations soon grow too broad to be displayable on a screen. The correspondence of natural deduction and sequent calculus

can be used for the translation of such sequent derivations. The most straightforward correspondence is the isomorphic translation between cut-free sequent derivations and normal natural derivations that use general elimination rules, a correspondence defined in von Plato (2001). The general rules are:

Table 3. General E-rules for $\&, \supset, \forall$.

$$
\cfrac{A \& B \qquad \overset{\overset{1\ 1}{A, B}}{\underset{\vdots}{C}}}{C} \ {\scriptstyle \& E, 1} \qquad\qquad \cfrac{A \supset B \quad A \quad \overset{\overset{1}{B}}{\underset{\vdots}{C}}}{C} \ {\scriptstyle \supset E, 1} \qquad\qquad \cfrac{\forall x A \qquad \overset{\overset{1}{A(t/x)}}{\underset{\vdots}{C}}}{C} \ {\scriptstyle \forall E, 1}
$$

These rules can be written in a linear notation exactly as rule $\exists E$. A schematic writing for rule $\& E$ is:

$$
\begin{array}{lll}
n. & A \& B & \\
& \vdots & \\
\hline
m. & A & \text{hyp } \& E, n \\
m+1. & B & \text{hyp } \& E, n \\
& \vdots & \\
k. & C & \\
\hline
k+1. & C & n, m\text{--}k, \& E
\end{array}
$$

Rule $\supset E$ is written similarly as:

$$
\begin{array}{lll}
n. & A \supset B & \\
& \vdots & \\
m. & A & \\
& \vdots & \\
\hline
k. & B & \text{hyp } \supset E, n \\
& \vdots & \\
l. & C & \\
\hline
l+1. & C & n, m, k\text{--}l, \supset E
\end{array}
$$

The rule for $\forall E$ is just like $\exists E$ except that there is no variable condition:

$n.$ $\forall x A(x)$

$\vdots$

$m.$	$A(t)$	hyp $\forall E, n$
	$\vdots$	
$k.$	C	

$k+1.$ C $n, m\text{–}k, \& E$

Cut-free derivation in an intuitionistic sequent calculus such as *G3i* can be displayed as a linear derivation through the translation to natural deduction followed by linearization with these rules.

References

[1] H. Curry, (1965) *Remarks on inferential deduction*, [in:] A. Tyminiecka (ed.), **Contributions to Logic and Methodology in Honor of J. M. Bochenski**, North-Holland (1965), pp. 45–72.

[2] G. Gentzen, *Untersuchungen über das logische Schliessen*, **Mathematische Zeitschrift**, vol. 39 (1934-35), pp. 176–210 and 405-431.

[3] S. Jaśkowski, *On the rules of supposition in formal logic* (1934), as reprinted [in:] S. McCall (ed.), **Polish Logic 1920–1939**, pp. 232–258, Oxford U. P. 1967.

[4] J. von Plato, *Natural deduction with general elimination rules*, **Archive for Mathematical Logic**, vol. 40 (2001), pp. 541–567.

[5] J. von Plato, **Elements of Logical Reasoning**, Cambridge, 2013.

[6] D. Prawitz, **ABC i Symbolisk Logik**, Mimeographed compendium. Later expanded and printed editions with Bokförlaget Thales, 1973.

Department of Philosophy
University of Helsinki
Helsinki, Finland
e-mail: jan.vonplato@helsinki.fi