Cut Elimination in a Category-like Sequent System

Zoran Petrić
Faculty of Mining and Geology
Djušina 7, 11000 Belgrade, Yugoslavia

Dedicated to Professor Slaviša Prešić

Abstract

A sequent system $\mathcal{L}$ for the conjunction-implication fragment of intuitionistic propositional logic is introduced. Sequents of $\mathcal{L}$ are of the form $A \vdash B$, where A and B are formulae, i.e. sequences of formulae with exactly one member. With a modification of Gentzen's procedure a cut elimination theorem for $\mathcal{L}$ is proved. Some categorial consequences of this result are pointed out.

1 Introduction

The work on this note was inspired by the paper of Kelly and MacLane [1971] where the cut elimination procedure was used to prove two facts connected with symmetric monoidal closed categories, namely the naturality of its canonical transformations and the property of coherence. The authors were inspired by Lambek (see [1968]) who was the first who has used a cut-elimination technique in category theory. However, we stay here in the logical framework and try to clarify the process of preparation of a logical system for further categorial purposes.

1.1 System $\mathcal{L}$

The sequent system $\mathcal{L}$ for the conjunction-implication fragment of intuitionistic propositional logic is introduced as follows. *Formulae* of the logic are built from an infinite set of *propositional letters* and the constant $\top$, by the logical connectives $\wedge$ and $\rightarrow$. The set of all formulae is denoted by $\mathcal{F}$. *Sequents* of $\mathcal{L}$ are of the form $A \vdash B$ for A and B in $\mathcal{F}$. We call A in $A \vdash B$ the *antecedent*, and B the *consequent* of the sequent. In order to introduce the rules of inference of $\mathcal{L}$ we need the following auxiliary notion of $\wedge$-*context*, which corresponds to the notion of (poly)functor in categories. A $\wedge$-context is defined inductively as follows:

1° The symbol $\Box$ is a $\wedge$-context.

2° If F is a $\wedge$-context and $A \in \mathcal{F}$, then $(F \wedge A)$ and $(A \wedge F)$ are $\wedge$-contexts.

3° If F and G are $\wedge$-contexts, then $(F \wedge G)$ is a $\wedge$-context.

For a $\wedge$-context F we say that it is a $\wedge_1$-context if the symbol $\Box$ occurs in F exactly once. For F a $\wedge$-context and $A \in \mathcal{F}$, we obtain $F(A)$ by substituting A for $\Box$ in F, e.g. if $F \equiv (B \wedge \Box) \wedge C$, then $F(A) = (B \wedge A) \wedge C$.

The *axioms* of $\mathcal{L}$ are

$$a_A : A \vdash A, \quad \text{for every} \quad A \in \mathcal{F},$$

The *structural rules* of $\mathcal{L}$ are

$$(\beta^{\leftarrow}_F)\ \frac{F(A \wedge (B \wedge C)) \vdash D}{F((A \wedge B) \wedge C) \vdash D} \qquad (\beta^{\rightarrow}_F)\ \frac{F((A \wedge B) \wedge C) \vdash D}{F(A \wedge (B \wedge C)) \vdash D}$$

$$(\gamma_F)\ \frac{F(A \wedge B) \vdash C}{F(B \wedge A) \vdash C}$$

$$(\omega_F)\ \frac{F(A \wedge A) \vdash B}{F(A) \vdash B} \qquad (\theta^A_F)\ \frac{F(\top) \vdash B}{F(A) \vdash B}$$

$$(\tau_F)\ \frac{F(A) \vdash B}{F(A \wedge \top) \vdash B} \qquad (\tau'_F)\ \frac{F(A \wedge \top) \vdash B}{F(A) \vdash B}$$

$$(\mu_G)\ \frac{A \vdash B \qquad G(B) \vdash C}{G(A) \vdash C},$$

where F is a $\wedge_1$ context and G is a $\wedge$ context.

The *rules for connectives* are

$$(\wedge)\ \frac{A \vdash C \qquad B \vdash D}{A \wedge B \vdash C \wedge D}$$

$$(*)\ \frac{A \wedge B \vdash C}{B \vdash A \rightarrow C} \qquad (\triangle)\ \frac{A \vdash B \qquad C \wedge D \vdash E}{(A \wedge (B \rightarrow C)) \wedge D \vdash E}$$

A *proof* of a sequent $A \vdash B$ in $\mathcal{L}$ is a binary tree with sequents in its nodes, such that $A \vdash B$ is in the root, axioms are in the leaves and consecutive nodes are connected by some of the inference rules above.

What are the differences between $\mathcal{L}$ and the corresponding fragment of Gentzen's system LJ (see [1935])? In $\mathcal{L}$ we have just one meta-logical symbol ($\vdash$) in the sequents and we omit Gentzen's commas in the antecedents, whose role is now covered by the logical connective $\wedge$. Also, we can't have empty either the antecedent or the consequent of a sequent in $\mathcal{L}$. The logical constant $\top$ serves to fill gaps in antecedents. These discrepancies between $\mathcal{L}$ and LJ arise because in $\mathcal{L}$ we want antecedents and consequents of sequents to be of the same sort (namely members of $\mathcal{F}$) and this enables us to look at an $\mathcal{L}$ sequent as an arrow with the source being the antecedent and the target the consequent of the sequent.

Also, the rule ($\wedge$) is a rule of simultaneous introduction of the connective $\wedge$ on the both sides of a sequent: there is no a counterpart for this rule in LJ. This difference is not categorially motivated. We believe that $\mathcal{L}$ completely separates structural rules from the rules for connectives. On the other hand, the LJ rules &-IS and &-IA (see 1.22. of [1935]) have hidden interchanges, contractions and thinnings.

Since we prove the cut-elimination theorem through elimination of *mix*, as Gentzen did too, we have postulated mix rule (μ) as primitive. However this mix is something different from Gentzen's mix. It is liberal in the sense that the $\wedge$-context G in (μ_G) needn't to capture all factors B (see the definition below) as arguments in $G(B)$. The formula B may also be used in Step 2° of the construction of the $\wedge$-context G, i.e. mix needn't to "swallow" all the occurrences of B in $G(B)$. Also, there are no categorial reasons to prefer cut to such a mix. In both cases, we don't have categorial composition of arrows corresponding to both premises of the rule, but a more involved composition of the right premise with an image of the left premise under the functor corresponding to a $\wedge$-context. The only difference is that in the case of cut this is always a $\wedge_1$-context.

An advantage of $\mathcal{L}$ is that its proofs can be easily coded. For example the proof

$$\dfrac{p \vdash p \qquad \dfrac{q \vdash q}{q \wedge \top \vdash q}}{\dfrac{(p \wedge (p \to q)) \wedge \top \vdash q}{p \wedge (p \to q) \vdash q}}$$

is coded by

$$\tau'_{\Box}(a_p \triangle \tau_{\Box} a_q)$$

This fact helps when we want to postulate equalities that should hold between the proofs of $\mathcal{L}$.

For the proof of our main result we need the following notions of *degree* and *rank*. The degree of a formula is the number of logical connectives in it. However, because of the categorially motivated elimination of the comma, the symbol $\wedge$ plays a double role and in order to define rank, we define as follows a set of *factors* of A, for every $A \in \mathcal{F}$:

1° A is a factor of A,

2° if A is of the form $A_1 \wedge A_2$ then every factor of A_1 or A_2 is a factor of A.

Now, we introduce (in the style of Došen) an auxiliary indexing of consequents and factors of antecedents in a mixless proof of $\mathcal{L}$ which will help us in defining the rank of an occurrence of a formula in such a proof. First we index all the consequents and all the factors of antecedents of axioms by 1 and inductively proceed as follows. In all the structural rules and the rule $(\triangle)$ the index of the consequent in the conclusion is increased by 1. In $(\wedge)$ and $(*)$ the index of the consequent in the conclusion is 1. Every factor of the antecedent preserved by a rule has the index increased by 1, and all the factors introduced or modified by the rule (take care that we always speak about occurrences of formulae and not just about formulae) have index 1 in the conclusion. In (ω_F) the occurrence of A in the conclusion is indexed by the maximum of indices of distinguished A's in the premise, increased by 1. In the example of the proof given above this indexing looks like

$$\dfrac{p^1 \vdash p^1 \qquad \dfrac{q^1 \vdash q^1}{(q^2 \wedge \top^1)^1 \vdash q^2}}{\dfrac{((p^2 \wedge (p \to q)^1)^1 \wedge \top^2)^1 \vdash q^3}{(p^3 \wedge (p \to q)^2)^2 \vdash q^4}}$$

Then the rank of an occurrence of a formula in a proof is given by its index.

2 Cut-elimination theorem and consequences

Our main result is the following.

Theorem *Every proof in $\mathcal{L}$ can be transformed into a proof of the same root-sequent with no applications of the rule (μ).*

Proof: As in the standard cut-elimination procedure it is enough to consider a proof whose last rule is (μ) and there is no more applications of (μ) in the proof. So let our proof be of the form

$$\dfrac{\begin{array}{c}\pi_1\\ A \vdash B\end{array} \qquad \begin{array}{c}\pi_2\\ G(B) \vdash C\end{array}}{G(A) \vdash C}$$

with π_1 and π_2 mixless. Then we define the *degree of this proof* as the degree of B and the *rank of this proof* as the sum of the *left rank*, i.e. the rank of the occurrence of B in the left premise of

(μ), in the subproof π_1, and the *right rank*, i.e. the maximum of all ranks of distinguished factors B in the right premise of (μ) in the subproof π_2. Then we prove our theorem by induction on the lexicographically ordered pairs $\langle d, r\rangle$ for the degree d and the rank r of the proof.

1. $r = 2$

The following situations should be considered: 1.1. π_1 or π_2 are axioms; 1.2. π_1 ends with ($\wedge$); 1.3.1. π_1 ends with ($*$) and π_2 ends with ($\triangle$); 1.3.2. π_1 ends with ($*$) and π_2 ends with (θ^B). We illustrate here just Case 1.2.

Suppose our proof is of the form

$$\dfrac{\dfrac{\begin{matrix}\pi_1' \\ A_1 \vdash B_1\end{matrix} \qquad \begin{matrix}\pi_1'' \\ A_2 \vdash B_2\end{matrix}}{A_1 \wedge A_2 \vdash B_1 \wedge B_2}\,\wedge \qquad \begin{matrix}\pi_2 \\ G(B_1 \wedge B_2) \vdash C\end{matrix}}{G(A_1 \wedge A_2) \vdash C}\,\mu$$

Then this proof is transformed into the proof

$$\dfrac{\begin{matrix}\pi_1' \\ A_1 \vdash B_1\end{matrix} \qquad \dfrac{\begin{matrix}\pi_1'' \\ A_2 \vdash B_2\end{matrix} \qquad \begin{matrix}\pi_2 \\ G(B_1 \wedge B_2) \vdash C\end{matrix}}{G(B_1 \wedge A_2) \vdash C}\,\mu}{G(A_1 \wedge A_2) \vdash C}\,\mu$$

where both applications of (μ) have lower degree.

2. $r > 2$

The following situations should be considered: 2.1. π_2 ends with a structural rule; 2.2. π_2 ends with ($\wedge$); 2.3. π_2 ends with ($*$); 2.4. π_2 ends with ($\triangle$); 2.5. π_1 ends with a structural rule; 2.6. π_1 ends with ($\triangle$). Cases 2.1. - 2.4. are considered under the assumption that the right rank is greater than 1, while 2.5. and 2.6. are connected with the assumption that the left rank is greater than 1. Case 2.1. includes a lot of subcases and we illustrate one of them here.

Suppose our proof is of the form

$$\dfrac{\begin{matrix}\pi_1 \\ A \vdash B\end{matrix} \qquad \dfrac{\begin{matrix}\pi_2 \\ G_1(B) \vdash C\end{matrix}}{G(B) \vdash C}\,\beta^{\rightarrow}}{G(A) \vdash C}\,\mu$$

where G is obtained from G_1 by substituting $H \wedge \Box$ for a subcontext $(H \wedge B_1) \wedge (B_2 \wedge B_3)$ of G_1, and H is a $\wedge$-context and $B \equiv B_1 \wedge (B_2 \wedge B_3)$. We call this new box of G the *principal box*. Then this proof is transformed into the proof

$$\dfrac{\begin{matrix}\pi_1 \\ A \vdash B\end{matrix} \qquad \dfrac{\dfrac{\begin{matrix}\pi_1 \\ A \vdash B\end{matrix} \qquad \begin{matrix}\pi_2 \\ G_1(B) \vdash C\end{matrix}}{G_1(A) \vdash C}\,\mu}{G_2(B) \vdash C}\,\beta^{\rightarrow}}{G(A) \vdash C}\,\mu$$

where G_2 is obtained from G by substituting A for all boxes except the principal one which remains the unique box in G_2. Then the upper application of (μ) has its rank decreased by one and the right rank of the lower application of (μ) is 1. $\square$

It is possible to check that all the reduction steps of our cut-elimination procedure are covered by the equalities of cartesian closed categories which can be naturally defined in the language of $\mathcal{L}$. These equations are sufficient for cut elimination, but they need not all be necessary. This is an argument for the justification of these categories. However, the main consequence of our Theorem is another proof of the result from [1992], which claims that all canonical transformations from cartesian closed categories are natural in the extended sense. The fact that all proofs of $\mathcal{L}$ can be reduced to a cut-free form, directly eliminates all obstacles in the way of naturality. This result was originally proved by the apparatus of natural deduction, and this is an alternative, sequent system, approach.

References

Gentzen, G. [1935] Untersuchungen über das logische Schließen. *Mathematische Zeitschrift* **39** 176-210, 405-431 (English translation in *The Collected Papers of Gerhard Gentzen*, Szabo, M.E. (ed.), North-Holland).

Girard, J.-Y., Scedrov, A., and Scott, P.J. [1992] Normal forms and cut-free proofs as natural transformations. In: Moschovakis, Y.N. (ed.) *Logic from Computer Science, Proceedings of a workshop held November 13-17, 1989*, Springer, 217-241.

Kelly, G.M., and Mac Lane, S. [1971] Coherence in closed categories, *Journal of Pure and Applied Algebra* **1**, 97-140.

Lambek, J. [1968] Deductive systems and categories I: Syntactic calculus and residuated categories, *Mathematical Systems Theory* **2**, 287-318.