A Note on the Use of Sum in the Logic of Proofs

Roman Kuznets*

Institut für Informatik und angewandte Mathematik
Universität Bern
Neubrückstrasse 10, 3012 Bern, Switzerland
kuznets@iam.unibe.ch

Abstract. The Logic of Proofs LP, introduced by Artemov, encodes the same reasoning as the modal logic S4 using proofs explicitly present in the language. In particular, Artemov showed that three operations on proofs (application $\cdot$, positive introspection !, and sum $+$) are sufficient to mimic provability concealed in S4 modality. While the first two operations go back to Gödel, the exact role of $+$ remained somewhat unclear. In particular, it was not known whether the other two operations are sufficient by themselves. We provide a positive answer to this question under a very weak restriction on the axiomatization of LP.

The Logic of Proofs LP was introduced in [Art95] as an explicit counterpart of the modal logic of provability S4 (an almost identical format was suggested by Gödel in a lecture that remained unpublished until [Göd95]). In this new format, the inability to interpret the Reflection Principle directly into formal arithmetic is overcome by an introduction of explicit proofs into the language. These proof objects, with some basic operations on them, are sufficient to fully realize all valid facts about provability that can be formulated in a less precise modal language. These proof objects, which Artemov called *proof polynomials* or *proof terms*, are constructed according to the following grammar:

$$t ::= x \mid c \mid (t \cdot t) \mid (t + t) \mid !\, t \ ,$$

where x stands for a proof variable, c stands for a proof constant, and the operations $+$ and $\cdot$ are by default associated to the left. Proof polynomials enable us to build new formulas via a new formula construct $t{:}F$ read *t is a proof of F*. The language of LP is obtained by adding this construct to the propositional language, e.g., with $\bot$ and $\to$ as basic connectives (by default, we will consider $\to$ to be associated to the right).

In fulfillment of Gödel's original program, in [Art95] (see also [Art01]) it was shown that the Logic of Proofs LP in this language serves as a missing link between S4 and provability in formal arithmetic, thereby providing a long-sought provability semantics for S4. The axioms and rules of the Logic of Proofs are as follows:

* Supported by Swiss National Science Foundation grant 200021–117699.

Axioms and rules of LP:

A1. A complete axiomatization of classical propositional logic by finitely many axiom schemes; rule modus ponens;

A2. *Application Axiom* $s{:}(F \to G) \to (t{:}F \to (s \cdot t){:}G)$;

A3. *Sum Axiom* $s{:}F \to (s+t){:}F, \quad t{:}F \to (s+t){:}F$;

A4. *Reflection Axiom* $t{:}F \to F$;

A5. *Proof Checker Axiom* $t{:}F \to {!}t{:}t{:}F$;

R4. *Axiom Internalization Rule*: $\dfrac{}{c{:}A}$, where A is an axiom and c is a proof constant.

The connection between S4 and formal arithmetic via LP is given by the following two statements through an operation $(\cdot)^\circ$ of *forgetful projection*, defined recursively by $p^\circ = p$, $\bot^\circ = \bot$, $(F \to G)^\circ = F^\circ \to G^\circ$, $(t{:}F)^\circ = \Box(F^\circ)$:

Theorem 1 (Arithmetical Completeness, [Art95, Art01]). *For any* LP*-formula* F, $\mathsf{LP} \vdash F$ *iff all arithmetical interpretations*[1] *of* F *are provably valid.*

Theorem 2 (Realization Theorem, [Art95, Art01]). *For any modal formula* F, $\mathsf{S4} \vdash F$ *iff* $\mathsf{LP} \vdash F^r$ *for some* LP*-formula* F^r *such that* $(F^r)^\circ = F$.

Thus, a modal statement about provability can first be refined by realizing occurrences of $\Box$ with specific proof polynomials; the resulting LP-formula can then be interpreted in the arithmetical language. The proof of the Realization Theorem is constructive: there exists an algorithm that restores proof polynomials to replace $\Box$'s in any given modal theorem. More precisely, the algorithm requires a cut-free Gentzen derivation of the given modal formula as input and constructs polynomials based on this derivation. The following lemma is a fundamental feature of proof terms and is central to the realization algorithm:

Lemma 1 (Lifting Lemma, [Art95, Art01]). *If*

$$G_1, \ldots, G_n, s_1{:}H_1, \ldots, s_k{:}H_k \vdash_{\mathsf{LP}} F \ ,$$

it is possible to construct a $+$*-free proof term* $t(x_1, \ldots, x_n, y_1, \ldots, y_k)$ *such that*

$$x_1{:}G_1, \ldots, x_n{:}G_n, s_1{:}H_1, \ldots, s_k{:}H_k \vdash_{\mathsf{LP}} t(x_1, \ldots, x_n, s_1, \ldots, s_k){:}F \ .$$

Thus, all derivations in LP can be emulated by operations on proof terms without the use of $+$. The other two operations on proofs—application $\cdot$ and proof checker $!$—were present already in Gödel's lecture [Göd95] (in a slightly different but essentially equivalent form). The former is the internalized version of modus ponens, and the latter is a proof verifier. Note that the forgetful projections of the corresponding axioms, A2 and A5 respectively, are nothing but familiar Hilbert–Bernays–Löb postulates about provability.

The operation $+$ in Artemov's system was a novelty. It corresponds to the combination of two proofs, e.g., concatenation of Hilbert-style derivations where

[1] We are omitting all technical details of what an arithmetical interpretation is since this is not the focus of this paper; these details can be found in [Art01].

the resulting derivation proves all the formulas proven by either of the concatenated derivations. This operation firmly plants LP in the realm of multi-conclusion proofs, i.e., proof terms must be capable of proving multiple formulas at the same time.[2] But it is not immediately clear how crucial the presence of $+$ is for the Realization Theorem, although $+$ is actively used in Artemov's realization algorithm, as well as in other realization proofs (see [Art95, Art01, Fit05, BK06, Fit07]). Fitting showed in [Fit05] that all S4-theorems can be realized without $+$ if subformulas in them are allowed to be duplicated. But $+$ was still used to eliminate the duplicates. In this paper, we show that S4 can actually be realized in LP without $+$ although proof terms must remain multi-conclusion.

Artemov's realization algorithm from [Art95] works by induction on a given cut-free Gentzen derivation of a given S4-theorem. We will briefly recapture the $+$-sensitive steps of the algorithm referring the reader to the source for further details. All occurrences of $\Box$ (or simply $\Box$'s for brevity) in the derivation are broken into families of related ones. Clearly, all $\Box$'s from a family must be realized by the same proof term. Note that in any cut-free derivation, the polarity of all $\Box$'s from one family must be the same. Negative families can be realized by arbitrary proof terms (e.g., by *distinct* proof variables in the original algorithm), and so can positive families if no $\Box$ in them is introduced by a $(\Box R)$-rule. For our purposes, it is important to realize all negative families by the *same* proof variable x. The algorithm realizes each sequent $\Gamma \Rightarrow \Delta$ in the derivation by an LP-theorem $t:(\wedge\Gamma^r \to \vee\Delta^r)$ by induction on the depth of the Gentzen derivation, with the *outer term* t capturing a Hilbert derivation of the realization $\wedge\Gamma^r \to \vee\Delta^r$ of the sequent $\Gamma \Rightarrow \Delta$.

The only part of the algorithm that requires the use of $+$ is the realization of the positive $\Box$'s in front of F in $(\Box R)$-rules, which have the form

$$\frac{\Box\Gamma \Rightarrow F}{\Box\Gamma \Rightarrow \Box F}\quad (\Box R)\ . \tag{1}$$

For lack of space, we only outline the differences from Artemov's way of realizing the $\Box$ in front of F. By the Lifting Lemma, the Deduction Theorem, and propositional reasoning, there is a term $t(x)$ such that $\mathsf{LP} \vdash x:\Gamma^r \to t(x):(x:\Gamma^r)$, where x is the proof variable that realizes all negative $\Box$'s, $\Gamma = \{G_1,\dots,G_n\}$ is a multiset, $\Gamma^r = \{G_1^r,\dots,G_n^r\}$ is its realization, and $x{:}\Gamma^r = x{:}G_1^r \wedge \dots \wedge x{:}G_n^r$ realizes the common antecedent of both premise and conclusion of (1).

In general, the ith introduction of $\Box F$ from a given family by $(\Box R)$ yields its own term $t_i(x)$. The outer terms s_i for the premises of $(\Box R)$-rules such that, by IH, $\mathsf{LP} \vdash s_i:(x:\Gamma_i^r \to F^r)$, where F^r is the realization of F, also differ. Artemov's algorithm combined different realizations $s_i \cdot t_i(x)$ of the positive $\Box$ in front of F using plus: namely, Artemov showed that $\mathsf{LP} \vdash x:\Gamma_i^r \to q:F^r$, $i = 1,\dots,N$, for $q = s_1 \cdot t_1(x) + \dots + s_N \cdot t_N(x)$, where N is the number of $(\Box R)$-rules used to introduce $\Box$'s from the family. Thus, this q can serve as a realization for this family of $\Box$'s (outer terms for the conclusions of the rules are then easy to construct using the Lifting Lemma).

[2] There exist single-conclusion versions of LP (see [Kru01]), but they cannot correspond to any normal modal logic.

If $\Gamma_i = \emptyset$ for some $1 \le i \le N$, i.e., if some of the $(\Box R)$-rules in the family infer $\Rightarrow \Box F$ from $\Rightarrow F$, no $+$ is necessary for realizing the $\Box$ in front of F. By IH, in this case, $\mathsf{LP} \vdash F^r$. By the Lifting Lemma, there exists a ground $+$-free term q' such that $\mathsf{LP} \vdash q' : F^r$. Clearly, this term can be used to realize the whole family because in this case $\mathsf{LP} \vdash x : \Gamma_i^r \to q' : F^r$ for all $i = 1, \dots, N$.

So let us assume that $\Gamma_i \neq \emptyset$ for all $1 \le i \le N$. To be able to avoid the use of $+$, we need to unify all $t_i(x)$ from the same family. We will use the following

Definition. Each formula is a *balanced conjunction (disjunction) of depth* 0. If A and B are both balanced conjunctions (disjunctions) of depth k, then $A \wedge B$ $(A \vee B)$ is a *balanced conjunction (disjunction) of depth* $k+1$.[3]

Note that the number of conjuncts (disjuncts) in a balanced conjunction (disjunction) is always a degree of 2. We will require that all $x : \Gamma_i^r$, $i = 1, \dots, N$, be fully balanced conjunctions of the same depth. Making them balanced is not difficult if they all have the same length, a degree of 2. Hence, our task is to inflate them artificially to the same size without changing the meaning of formulas $x : \Gamma_i^r$. This can be achieved by duplicating the last element in each Γ_i sufficiently many times (remember that $\Gamma_i \neq \emptyset$), using Weakening, and associating the resulting conjunction in a balanced way. These duplicate assumptions are eliminated by Contraction immediately after the $(\Box R)$-rule. Neither Weakening nor Contraction present difficulties for a $+$-free realization.

It is easy to check that for any collection of sets Γ_i^r, each of cardinality 2^k, and for corresponding balanced conjunctions $x : \Gamma_i^r$, $\mathsf{LP} \vdash x : \Gamma_i^r \to t^k(x) : (x : \Gamma_i^r)$ for terms $t^k(x)$ recursively defined by $t^0(x) = !\, x$, $t^{k+1}(x) = c_\wedge \cdot t^k(x) \cdot t^k(x)$, where $\mathsf{LP} \vdash c_\wedge : (A \to B \to A \wedge B)$ for any LP-formulas A and B. The existence of such $+$-free $c_\wedge$ is guaranteed by the Lifting Lemma and the Substitution Property.

We now show how to replace Artemov's term q for a positive family of $\Box$'s with a $+$-free term q' such that $\mathsf{LP} \vdash x : \Gamma_i^r \to q' : F^r$, $i = 1, \dots, N$. The rest of Artemov's original algorithm remains unchanged and effectively yields a $+$-free realization. We use

$$q' = \mathrm{syl}\Big(d_n,\ e_n \cdot s_1 \cdots s_{2^n}\Big) \cdot t^k(x)\ , \tag{2}$$

where $2^{n-1} < N \le 2^n$, $s_{N+1} = \ldots = s_{2^n} = s_N$, the cardinality of all Γ_i^r's is 2^k,

$$\mathsf{LP} \vdash e_n : \Big((A_1 \to B) \to \cdots \to (A_{2^n} \to B) \to (A_1 \vee \cdots \vee A_{2^n} \to B)\Big)\ , \tag{3}$$

$$t : (A \to B), s : (B \to C) \quad \vdash_{\mathsf{LP}} \quad \mathrm{syl}(t, s) : (A \to C)\ , \tag{4}$$

$$\mathsf{LP} \vdash d_n : (A_i \to A_1 \vee \cdots \vee A_{2^n}), \quad i = 1, \dots, 2^n \tag{5}$$

for any formulas $A_1, \dots, A_{2^n}$, A, B, and C and any terms t and s, where disjunctions $A_1 \vee \cdots \vee A_{2^n}$ are always balanced of depth n. The existence of the $+$-free term e_n from (3) follows from the Lifting Lemma and the Substitution Property. So does the existence of $+$-free $\mathrm{syl}(x, y)$ from (4); it is explicitly constructed in [BK09]. It is not hard to show that d_n can be recursively defined

[3] Balanced conjunctions were first applied in the context of LP in [BK09].

by making $d_0 = x$, $d_{n+1} = \mathrm{syl}(d_n, c_\vee)$, where $\mathsf{LP} \vdash c_\vee : (A \to A \vee B)$ and $\mathsf{LP} \vdash c_\vee : (B \to A \vee B)$ for any formulas A and B. The existence of such a $+$-free term $c_\vee$ is not generally guaranteed. But if axiom schemes $A \to A \vee B$ and $B \to A \vee B$ are both present in A1, then any constant can serve as $c_\vee$. We have proved the following

Theorem 3. *If $A \to A \vee B$ and $B \to A \vee B$ are axioms of* LP *for any formulas A and B, then modal logic* S4 *can be realized in* LP *without the use of $+$.*

It should be noted that Artemov's original Realization Theorem put an additional restriction of *normality* on realizations, namely, that all negative $\Box$'s in an S4-theorem should be realized by distinct proof variables. The $+$-free realization constructed in Theorem 3 is not normal because all negative $\Box$'s in it are realized by the same proof variable rather than by many distinct ones. The proof of the following theorem is omitted due to space constraints.

Theorem 4. *There are theorems of modal logic* S4 *that do not have normal $+$-free realizations in* LP.

Acknowledgments.

We thank Remo Goetschi for discussions leading to an improvement of the result. We are grateful to Galina Savukova for proofreading this paper.

References

[Art95] Sergei N. Artemov. Operational modal logic. Technical Report MSI 95–29, Cornell University, December 1995.

[Art01] Sergei N. Artemov. Explicit provability and constructive semantics. *Bulletin of Symbolic Logic*, 7(1):1–36, March 2001.

[BK06] Vladimir Brezhnev and Roman Kuznets. Making knowledge explicit: How hard it is. *Theoretical Computer Science*, 357(1–3):23–34, July 2006. doi:10.1016/j.tcs.2006.03.010.

[BK09] Samuel R. Buss and Roman Kuznets. The NP-completeness of reflected fragments of justification logics. In Sergei [N.] Artemov and Anil Nerode, editors, *Logical Foundations of Computer Science, International Symposium, LFCS 2009, Deerfield Beach, FL, USA, January 3–6, 2009, Proceedings*, volume 5407 of *Lecture Notes in Computer Science*, pages 122–136. Springer, 2009. doi:10.1007/978-3-540-92687-0_9.

[Fit05] Melvin Fitting. The logic of proofs, semantically. *Annals of Pure and Applied Logic*, 132(1):1–25, February 2005. doi:10.1016/j.apal.2004.04.009.

[Fit07] Melvin Fitting. Realizations and LP. In Sergei N. Artemov and Anil Nerode, editors, *Logical Foundations of Computer Science, International Symposium, LFCS 2007, New York, NY, USA, June 4–7, 2007, Proceedings*, volume 4514 of *Lecture Notes in Computer Science*, pages 212–223. Springer, 2007. doi:10.1007/978-3-540-72734-7_15.

[Göd95] Kurt Gödel. Vortrag bei Zilsel/Lecture at Zilsel's (*1938a). In Solomon Feferman, John W. Dawson, Jr., Warren Goldfarb, Charles Parsons, and Robert M. Solovay, editors, *Unpublished essays and lectures*, volume III of *Kurt Gödel Collected Works*, pages 86–113. Oxford University Press, 1995.

[Kru01] Vladimir N. Krupski. The single-conclusion proof logic and inference rules specification. *Annals of Pure and Applied Logic*, 113(1–3):181–206, December 2001. doi:10.1016/S0168-0072(01)00058-6.