Calibrating Provability Logic: From Modal Logic to Reflection Calculus

Lev Beklemishev [1]

Steklov Institute of Mathematics, Moscow

Several interesting applications of provability logic in proof theory made use of a polymodal logic **GLP** due to Giorgi Japaridze. This system, although decidable, is not very easy to handle. In particular, it is not Kripke complete. It is complete w.r.t. neighborhood semantics, however this could only be established recently by rather complicated techniques [1].

In this talk we will advocate the use of a weaker system, called *Reflection Calculus*, which is much simpler than **GLP**, yet expressive enough to regain its main proof-theoretic applications, and more. From the point of view of modal logic, **RC** can be seen as a fragment of polymodal logic consisting of implications of the form $A \to B$, where A and B are formulas built-up from $\top$ and the variables using just $\wedge$ and the diamond modalities. In this paper we formulate it in a somewhat more succinct self-contained format.

Further, we state its arithmetical interpretation, and provide some evidence that **RC** is much simpler than **GLP**. We then outline a consistency proof for Peano arithmetic based on **RC** and state a simple combinatorial statement, the so-called Worm principle, that was suggested by the use of **GLP** but is even more directly related to the Reflection Calculus.

1 Reflection calculus RC

Basic symbols of **RC** are propositional variables $p, q, \ldots$, constant $\top$, conjunction $\wedge$, the symbols n, for each $n \in \omega$, and the brackets. Informally, n corresponds to the n-th modality $\langle n \rangle$.

The formulas α of **RC** are generated by the following grammar:

$$\alpha ::= \top \mid p \mid (\alpha \wedge \alpha) \mid n\alpha \qquad n \in \omega.$$

Example: $\alpha = 3(2\top \wedge 32\top)$. The symbols $\top$ occurring after a number symbol can be omitted without impairing the readability of the formula, e.g., the previous formula can be shortened to $3(2 \wedge 32)$.

Derivable objects of **RC** are *sequents*, that is, expressions of the form $\alpha \vdash \beta$ with α, β formulas.

[1] Supported by the Russian Foundation for Basic Research (RFBR), Russian Presidential Council for Support of Leading Scientific Schools, and the Swiss–Russian cooperation project STCP–CH–RU "Computational proof theory."

RC rules:

(i) $\alpha \vdash \alpha$; $\alpha \vdash \top$; if $\alpha \vdash \beta$ and $\beta \vdash \gamma$ then $\alpha \vdash \gamma$;

(ii) $\alpha \wedge \beta \vdash \alpha, \beta$; if $\alpha \vdash \beta$ and $\alpha \vdash \gamma$ then $\alpha \vdash \beta \wedge \gamma$;

(iii) $nn\alpha \vdash n\alpha$; if $\alpha \vdash \beta$ then $n\alpha \vdash n\beta$;

(iv) $n\alpha \vdash m\alpha$ for $n > m$;

(v) $n\alpha \wedge m\beta \vdash n(\alpha \wedge m\beta)$ for $n > m$.

For example, the following is derivable in **RC**:

$$3 \wedge 23 \vdash 3(\top \wedge 23) \vdash 323.$$

Notice that Axioms 1 and 2 express that $\vdash$ induces a Tarskian consequence relation and that $\wedge$ has the usual properties of conjunction. Axioms 3 correspond to the modal axioms of **K4**. (Notably, any principle related to Löb's axiom is absent.) Axioms 4 and 5 relate different modalities to each other.

In the following, the variable-free fragment of **RC** will, in a sense, be more important than **RC** itself. We denote it $\mathbf{RC}^0$.

2 Arithmetical interpretation of RC

Let S be a first order r.e. theory containing enough arithmetic to satisfy the assumptions of Gödel's second incompleteness theorem. For each $n \in \omega$, *reflection principles* $R_n(S)$ are the formulas in the language of S naturally expressing that *each arithmetical* Σ^0_n*-sentence provable in* S *is true*. Reflection principles are well-known in proof theory; their use is going back to Rosser, Turing, Kreisel and Feferman. They are best to be seen as generalizations of Gödel's consistency assertion to higher levels of arithmetical complexity.

Having fixed the formulas $R_n(S)$, we now define an interpretation of the language of **RC** in the style of provability logic.

Let f be a substitution mapping propositional variables to sentences in the language of S. *Arithmetical translation* $f_S(\alpha)$ of a formula α is defined inductively as follows:

- $f_S(\top) = \top$; $f_S(p) = f(p)$; $f_S(\alpha \wedge \beta) = (f_S(\alpha) \wedge f_S(\beta))$;
- $f_S(n\alpha) = R_n(S + f_S(\alpha))$.

Suppose $\mathbb{N} \vDash S$ and S contains Peano arithmetic PA.

Theorem 2.1 $\alpha \vdash \beta$ *in* **RC** *iff* $S \vdash f_S(\alpha) \to f_S(\beta)$*, for all* f.

We note that if α is variable-free, then $f_S(\alpha)$ does not depend on f. We abbreviate $f_S(\alpha)$ by α_S.

3 Interpretation of RC in GLP

As we mentioned before, **RC** can be seen as a fragment of polymodal provability logic **GLP**. We translate **RC**-formulas α to **GLP**-formulas α^* as follows: $\top^* = \top$, $p^* = p$, $(\alpha \wedge \beta)^* = (\alpha^* \wedge \beta^*)$, and $(n\alpha)^* = \langle n \rangle \alpha^*$. Thus, $3(2 \wedge 32)$ translates to $\langle 3 \rangle (\langle 2 \rangle \top \wedge \langle 3 \rangle \langle 2 \rangle \top)$.

The following theorem is an adaptation of the results of Dashkov [6].

Theorem 3.1 (i) **GLP** *is a conservative extension of* **RC**, *that is, for each* α, β, **RC** *proves* $\alpha \vdash \beta$ *iff* **GLP** $\vdash \alpha^* \to \beta^*$;

(ii) **RC** *is polytime decidable;*

(iii) **RC** *enjoys the finite model property.*

We note that by the results of Shapirovsky, **GLP** is PSPACE-complete. We also note that Theorem 2.1 follows from part 1 and Japaridze's arithmetical completeness theorem for **GLP**.

From now on we shall mainly work in the variable-free fragment of **RC**.

4 $\mathbf{RC}^0$ as an ordinal notation system

Let W denote the set of all $\mathbf{RC}^0$-formulas. Using derivability in **RC** we define the following relations on W:

- $\alpha \sim \beta$ if ($\alpha \vdash \beta$ and $\beta \vdash \alpha$);
- $\alpha <_n \beta$ if $\beta \vdash n\alpha$.

Obviously, $\sim$ is an equivalence relation and $<_n$ is correctly defined on the equivalence classes. We note that by the results of the previous section both of these relations are polynomially decidable.

A formula without variables and $\wedge$ is called a *word*. In fact, any such formula syntactically is a sequence of numbers (followed by $\top$).

Theorem 4.1 (i) *Every* $\alpha \in W$ *is equivalent to a word;*

(ii) $(W/\sim, <_0)$ *is isomorphic to* $(\varepsilon_0, <)$.

Here, ε_0 is the first ordinal α such that $\omega^\alpha = \alpha$. The isomorphism can be established by the following function $o : W/\sim \to \varepsilon_0$.

First, define $o(0^k) = k$, for each $k \in \omega$. Any other word can be written in the form $\alpha = \alpha_1 0 \alpha_2 0 \cdots 0 \alpha_n$, where each α_i does not contain 0 and not all of them are empty. Then we define

$$o(\alpha) = \omega^{o(\alpha_n^-)} + \cdots + \omega^{o(\alpha_1^-)},$$

where β^- means subtracting 1 from each letter of a word β.

Example 4.2 $o(1012) = \omega^{o(01)} + \omega^{o(0)} = \omega^{\omega^1 + \omega^0} + \omega = \omega^{\omega+1} + \omega$

Thus, calculating the ordinal $o(\alpha)$ gives a criterium for the equivalence and comparison of words. It is useful, however, to regard the set of words as a specific notation system for ordinals alternative to Cantor normal forms. In fact, in what follows we can completely disregard Cantor normal forms.

5 Reduction property

For the proof-theoretic applications of **RC** we need to state a basic property of reflection principles called the *reduction property*. Finitely iterated reflection

principles are defined as follows:

$$R_n^1(S) = R_n(S), \quad R_n^{k+1}(S) = R_n(S + R_n^k(S)).$$

Let A and B be two sets of formulas over a given arithmetical theory S. We write $A \equiv_n B$ modulo S, if $S + A$ and $S + B$ prove the same arithmetical Π^0_{n+1}-sentences. The following theorem is proved in [2].

Theorem 5.1 (reduction) *Suppose $S \subseteq \Pi^0_{n+2}$ and $V \vdash S$. Then*

$$R_{n+1}(V) \equiv_n \{R_n^k(V) : k < \omega\} \text{ modulo } S.$$

Let us now apply this theorem to the situation when $V = S + \beta_S$, for some $\beta \in W$.

Denote $\alpha = (n+1)\beta$ and $\alpha[\![0]\!] := n\beta$, $\alpha[\![k+1]\!] := n(\beta \wedge \alpha[\![k]\!])$.

It is easy to check that $\alpha[\![0]\!] <_0 \alpha[\![1]\!] <_0 \alpha[\![2]\!] <_0 \cdots \to \alpha$. Moreover, the formulas $\alpha[\![k]\!]$ correspond to k-fold iterated reflection principles $R_n^k(V)$. Thus, from the reduction property we infer

Corollary 5.2 *$\alpha_S \equiv_n \{\alpha[\![k]\!]_S : k < \omega\}$, whenever $S \subseteq \Pi^0_{n+2}$.*

6 Consistency proof for PA

Theorem 6.1 *Primitive recursive arithmetic together with transfinite induction over $(W, <_0)$ proves the consistency of* PA.

We sketch a proof of this version of Gentzen's theorem. As our basic system we take $S = \mathsf{EA}$, the Elementary Arithmetic, aka $I\Delta_0 + \exp$. We have that $\mathsf{EA} \subseteq \Pi^0_2$, so Corollary 5.2 applies for each n. We will also use the fact that PRA proves $R_1(\mathsf{EA})$.

Let $\Diamond\varphi$ denote a standard arithmetical formula expressing the consistency of a sentence (with the Gödel number) φ over S. In fact, $\Diamond\varphi$ is equivalent to $R_0(S + \varphi)$.

First, we prove $\forall\alpha\, \Diamond\alpha_S$ within PRA together with $(W, <_0)$-induction. Here and below, quantifiers $\forall\alpha$ are understood as ranging over Gödel numbers of words (under some natural Gödel numbering in S). Binary relation $<_0$ on W is arithmetized in a similar way.

It is sufficient to prove:

$$\mathsf{PRA} \vdash \forall\alpha\, (\forall\beta <_0 \alpha\, \Diamond\beta_S \to \Diamond\alpha_S).$$

The following argument can be formalized in PRA.

Assume $\forall\beta <_0 \alpha\, \Diamond\beta_S$.

- If $\alpha = 0\beta$, then $\Diamond\beta_S$. Since $\mathsf{PRA} \vdash R_1(S)$, every Π^0_1-sentence π implies $\Diamond\pi$. Taking $\Diamond\beta_S$ for π we infer $\Diamond\Diamond\beta_S$ and $\Diamond\alpha_S$.
- If $\alpha = (n+1)\beta$, then $\forall k\, \Diamond\alpha[\![k]\!]_S$, because $\alpha[\![k]\!] <_0 \alpha$.
  By Corollary 5.2 (formalizable in PRA),

$$\alpha_S \equiv_n \{\alpha[\![k]\!]_S : k < \omega\}.$$

Therefore $\forall k \Diamond\alpha[\![k]\!]_S$ yields $\Diamond\alpha_S$.

Thus, we have proved $\forall\alpha \Diamond\alpha_S$. What remains to be seen is that $\forall\alpha \Diamond\alpha_S$ implies the consistency of PA. This follows from a well known observation (originally due to Kreisel) that any instance of arithmetical induction follows from EA together with $R_n(\mathsf{EA})$, for an appropriate n. If, for each $n \in \omega$, the theory $S + n_S$ is consistent, then so is PA. In other words, $\forall n \Diamond n_S$ implies the consistency of PA.

7 The Worm principle

For any word α, we say that α *is higher than* n if each letter of α exceeds n. Given a word α, consider the following sequence $(\alpha_n)_{n\in\omega}$ of words.

Set $\alpha_0 := \alpha$ and suppose α_k is given. Define α_{k+1} by the following two rules:

- If $\alpha_k = 0\beta$ then $\alpha_{k+1} := \beta$.
- If $\alpha_k = (n+1)\beta$, find the longest (possibly empty) prefix β_0 of β such that β_0 is higher than n. Assume $\beta = \beta_0\gamma$. Then let $\alpha_{k+1} := (n\beta_0)^{k+2}\gamma$.

The *Worm principle* states that, for each α, the sequence α_k terminates in an empty word. A proof of the following theorem is based on the observation that the words α_k are equivalent to the formulas $\alpha[\![k]\!]$ in $\mathbf{RC}^0$ (see [4,3]).

Theorem 7.1 *The Worm principle is true but unprovable in Peano arithmetic. In fact, it is equivalent in* PRA *to the* Σ_1*-reflection* $R_1(\mathsf{PA})$ *for* PA.

The Worm principle can be seen as an analog of the well-known *Hydra battle* principle due to Paris and Kirby. However, it deals with words rather than finite trees. It can also be viewed, modulo some minor details, as a linear version of the so-called *Buchholz hydra battle* which deals with labeled trees. A version of the Worm principle deriving from Buchholz hydra battle has been analyzed by Hamano and Okada [7]. Independently but later, the Worm principle has been found (and baptized in the current form) in [4]. This paper was based on different, provability logical, considerations. A detailed correspondence between the Worm principle and the Hydra battle has been established by Carlucci [5], see also Lee [8].

References

[1] L. Beklemishev and D. Gabelaia. Topological completeness of the provability logic GLP. Preprint arXiv:1106.5693v1 [math.LO], 2011.

[2] L.D. Beklemishev. Provability algebras and proof-theoretic ordinals, I. *Annals of Pure and Applied Logic*, 128:103–123, 2004.

[3] L.D. Beklemishev. Reflection principles and provability algebras in formal arithmetic. *Uspekhi Matematicheskikh Nauk*, 60(2):3–78, 2005. In Russian. English translation in: *Russian Mathematical Surveys*, 60(2): 197–268, 2005.

[4] L.D. Beklemishev. The Worm Principle. In Z. Chatzidakis, P. Koepke, and W. Pohlers, editors, *Lecture Notes in Logic 27. Logic Colloquium '02*, pages 75–95. AK Peters, 2006. Preprint: Logic Group Preprint Series 219, Utrecht University, March 2003.

[5] L. Carlucci. Worms, gaps and hydras. *Mathematical Logic Quarterly*, 51(4):342–350, 2005.
[6] E.V. Dashkov. On a positive fragment of polymodal provability logic GLP. *Matematicheskie Zametki*, 91(3):331–336, 2012. English translation in: *Mathematical Notes* 91(3):318–333, 2012.
[7] M. Hamano and M. Okada. A relationship among Gentzen's proof-reduction, Kirbi-Paris' Hydra game, and Buchholz's Hydra game. *Mathematical Logic Quarterly*, 43(1):103–120, 1997.
[8] Gyesik Lee. A comparison of well-known ordinal notation systems for ε_0. *Annals of Pure and Applied Logic*, 147(1-2):48–70, 2007.