On Interpretability in the Theory of Concatenation

Vítězslav Švejdar*†

December 10, 2007

The original publication is available at Notre Dame J. of Formal Logic.

Abstract We prove that a variant of Robinson arithmetic Q with non-total operations is interpretable in the theory of concatenation TC introduced by A. Grzegorczyk. Since Q is known to be interpretable in that non-total variant, our result gives a positive answer to the problem whether Q is interpretable in TC. An immediate consequence is essential undecidability of TC.

1 Why weak theories, why concatenation?

Several versions of Gödel, Church, and Rosser theorems state the incompleteness and undecidability of every sufficiently strong recursively axiomatizable (consistent) theory T. The notion of "sufficiently strong" is usually made precise by stipulating that T extends Robinson arithmetic Q, or more generally, that T interprets Q. *Robinson arithmetic* Q, see [TMR53], is a theory useful from more than one point of view. It is finitely axiomatized and thus can be used in a straightforward proof of undecidability of first-order predicate logic. It is weak, but some richer arithmetics, like $\mathrm{I}\Delta_0$, are interpretable in it.

A natural question reads whether Q is the only or the best theory for explanation of incompleteness and undecidability phenomena. In connection with this question, A. Grzegorczyk in [Grz05] proposed to study the theory TC, the *theory of concatenation.* Instead of numbers that can be added and multiplied, in TC one has strings that can be concatenated, and there are two irreducible (single-letter) strings a and b. Some ideas behind formulation of axioms of TC go back to Quine [Qui46] and Tarski. Grzegorczyk's motivations to study the theory TC are philosophical and are explained in Introduction and in the beginning of Section 8 of [Grz05]. Speaking briefly, when reasoning, computing,

*This work is a part of the research plan MSM 0021620839 that is financed by the Ministry of Education of the Czech Republic. The author also acknowledges support by a Fulbright Scholarship at the Department of Philosophy, University of Notre Dame in Notre Dame, IN.

†Charles University, Prague, vitezslavdotsvejdaratcunidotcz, http://www1.cuni.cz/~svejdar/. Palachovo nám. 2, 116 38 Praha 1, Czech Republic.

or expressing knowledge, we deal with texts. Our ability to perform these tasks depends on discernibility, i.e. the possibility to identify and discriminate graphical objects. Thus decidability can be defined directly in terms of discernibility, without a reference to natural numbers. Then also the proof of undecidability of first order predicate logic could be more straightforward if based on strings and concatenation, because it can avoid coding of syntax, and also avoid the use of mathematical tools like Chinese Remainder Theorem.

The theory of concatenation is an interesting theory regardless whether one finds Grzegorczyk's motivations appealing. The paper [Grz05] contains a proof of undecidability of TC. Later Grzegorczyk and Zdanowski proved essential undecidability of TC in [GZ08], and left open the question whether Q is *interpretable* in TC. In the present paper we offer a positive answer to this question. Since a theory in which Q is interpretable must be essentially undecidable, see [TMR53], our result not only gives a piece of information missing in [GZ08]: it also yields an alternative proof of essential undecidability of TC.

A straightforward idea of constructing an interpretation of Q in TC is the following. Numbers are strings of the form $\mathtt{a}^k$, i.e. strings $\mathtt{a}^\frown\mathtt{a}^\frown\ldots^\frown\mathtt{a}$ where $\mathtt{a}$ is one of the two irreducible strings and $^\frown$ denotes concatenation. Addition of numbers is their concatenation. As to multiplication, $\mathtt{a}^k \cdot \mathtt{a}^n = \mathtt{a}^m$ if there exists a sequence w consisting of pairs such that (i) the first element of w is $[\mathtt{a}, \mathtt{a}^k]$, (ii) every element $[x, y]$ except the last one is immediately followed by an element $[x^\frown\mathtt{a}, y^\frown\mathtt{a}^k]$, and (iii) the last element is $[\mathtt{a}^n, \mathtt{a}^m]$. We basically follow this idea. However, there are some difficulties to be solved. For example, some expected properties of strings, like $\forall x\forall y\forall u(x^\frown u = y^\frown u \to x = y)$, are not provable in TC. It is difficult or even impossible to define in TC a reasonable notion of a sequence so that sequences of arbitrary lengths exist. And it is even not quite obvious how to define strings of the form $\mathtt{a}^k$ in TC. We show that none of these difficulties is essential. In particular, it does not matter that TC does not have sequences of arbitrary lengths, because what we really interpret in TC is a weaker variant of Q with possibly non-total addition and multiplication, not the full Q.

Thus we prove an existence of an interpretation of Q in TC by constructing an interpretation of a non-total variant of Q in TC and then combining this result with known facts that the interpretability relation is transitive and that Q is interpretable in that non-total variant. An obvious choice for the non-total variant is the theory Q^-, accidentally also introduced by A. Grzegorczyk, because a proof of interpretability of Q in Q^- can be found in [Šve07]. However, we will be able to advise a reader who wants to see a self-contained proof how to avoid reading the somewhat technically involved proof in [Šve07].

Another proof of interpretability of Q in TC, obtained independently but earlier than ours, is in Albert Visser's paper *Growing Commas* [Vis09]. While that paper is more general, the present paper was intended to be short and single-purpose, listing only those properties of TC needed for the main result.

Visser's paper, and also the paper [ČPR+07], contain also an information about *un*provability in TC and about its models. Yet other proofs of interpretability of Q in TC, independent of each other and of the present paper, were obtained by R. Sterken and M. Ganea [Gan09]. M. Ganea's proof is different from ours, but it also uses the result in [Šve07], i.e. uses the detour via Q^-.

2 Preliminaries: TC, Q^-, and the notion of interpretability

We work with somewhat different variant of the theory of concatenation than in [Grz05] and [GZ08], having an empty string ε and having three irreducible strings $\mathtt{a}$, $\mathtt{b}$, $\mathtt{c}$ rather than two. The exact choice of variant is inessential because all reasonable variants of TC are mutually interpretable ([GZ08, Vis09]). So our variant of the *theory of concatenation* TC has the language $\{^\frown, \varepsilon, \mathtt{a}, \mathtt{b}, \mathtt{c}\}$ with a binary function symbol and four constants. We systematically omit the symbol $^\frown$, i.e. write xy for the concatenation $x^\frown y$ of x and y. The axioms of TC are:

TC1: $\forall x(x\varepsilon = \varepsilon x = x)$,

TC2: $\forall x\forall y\forall z(x(yz) = (xy)z)$,

TC3: $\forall x\forall y\forall u\forall v(xy = uv \to \to \exists w((xw = u \ \&\ wv = y) \vee (uw = x \ \&\ wy = v)))$,

TC4: $\mathtt{a} \neq \varepsilon \ \&\ \forall x\forall y(xy = \mathtt{a} \to x = \varepsilon \vee y = \varepsilon)$,

TC5: $\mathtt{b} \neq \varepsilon \ \&\ \forall x\forall y(xy = \mathtt{b} \to x = \varepsilon \vee y = \varepsilon)$,

TC6: $\mathtt{c} \neq \varepsilon \ \&\ \forall x\forall y(xy = \mathtt{c} \to x = \varepsilon \vee y = \varepsilon)$,

TC7: $\mathtt{a} \neq \mathtt{b} \ \&\ \mathtt{a} \neq \mathtt{c} \ \&\ \mathtt{b} \neq \mathtt{c}$.

Our numbering of axioms of TC is more or less as in [Vis09], the difference is caused by the third letter $\mathtt{c}$ we have in the language. By axiom TC2, we can omit parentheses, and we do so almost everywhere. The axiom TC3 is called the *editor axiom* in [Grz05], and is attributed to Tarski. It describes what happens if two editors independently suggest splitting a large text into two volumes: the first volume of one of the editors consists of two parts, the other editor's first volume and a text (possibly empty) that appears as a starting part of the other editor's second volume.

The *theory* Q^-, weaker variant of Robinson arithmetic defined by A. Grzegorczyk, has language $\{0, \mathrm{S}, \mathrm{A}, \mathrm{M}\}$ with a constant, a unary function symbol, and two ternary relation symbols. The formulas $\mathrm{A}(x, y, z)$ and $\mathrm{M}(x, y, z)$ express that "z is the sum, or product, respectively, of x and y". The axioms of Q^- are:

A: $\forall x\forall y\forall z_1\forall z_2(\mathrm{A}(x, y, z_1) \ \&\ \mathrm{A}(x, y, z_2) \to z_1 = z_2)$,

M: $\forall x\forall y\forall z_1\forall z_2(\mathrm{M}(x, y, z_1) \ \&\ \mathrm{M}(x, y, z_2) \to z_1 = z_2)$,

Q1: $\forall x \forall y(\mathrm{S}(x) = \mathrm{S}(y) \rightarrow x = y)$,

Q2: $\forall x(\mathrm{S}(x) \neq 0)$,

Q3: $\forall x(x \neq 0 \rightarrow \exists y(x = \mathrm{S}(y)))$,

G4: $\forall x \mathrm{A}(x, 0, x)$,

G5: $\forall x \forall y \forall u(\exists z(\mathrm{A}(x, y, z) \ \& \ u = \mathrm{S}(z)) \rightarrow \mathrm{A}(x, \mathrm{S}(y), u))$,

G6: $\forall x \mathrm{M}(x, 0, 0)$,

G7: $\forall x \forall y \forall u(\exists z(\mathrm{M}(x, y, z) \ \& \ \mathrm{A}(z, x, u)) \rightarrow \mathrm{M}(x, \mathrm{S}(y), u))$.

Axioms Q1–Q3 are the same as in the full Robinson arithmetic Q, as defined in [TMR53]. Axioms G4–G7 are Grzegorczyk's reformulations of axioms Q4–Q7 of Q. They say that the number 0 can be added to any x from the right and that any x can be multiplied by 0 from the right, with the obvious results. If y can be added to x from the right then also $\mathrm{S}(y)$ can be added to x from the right. If x can be multiplied by y and the result is z, then it might not be possible to multiply x by $\mathrm{S}(y)$, which happens if the sum of z and x does not exist.

A *translation $*$ of formulas* of a theory T to formulas of a theory S is determined by a definitional extension S' of the theory S, a translation of symbols, and a domain. A *translation of symbols* maps each symbol of the theory T to a symbol of the definitional extension S' having the same kind (function or predicate) and arity. A *domain* is a formula $\delta(x)$ of S' with one free variable used to relativize quantifiers in the given translation $*$ of formulas: $(\forall x \varphi)^*$ is $\forall x(\delta(x) \rightarrow \varphi^*)$ and $(\exists x \varphi)^*$ is $\exists x(\delta(x) \ \& \ \varphi^*)$. The remaining logical symbols, i.e. connectives, are preserved by translation of formulas. One can think of the domain $\delta(x)$ as of the set $\{ \, x \, ; \, \delta(x) \, \}$, regardless whether the theory S comes with a notion of set. A translation $*$ of formulas is a (global, non-parametric, one-dimensional) *interpretation of T in S* if its domain $\delta(x)$ is (provably in S') non-empty and closed under all functions in the range of the corresponding translation of symbols, and if, in addition, $*$ maps all axioms of T to formulas provable in S'. A theory T is *interpretable in a theory S* if there exists an interpretation of T in S.

Interpretability can be taken as a measure of strength of axiomatic theories. If, for example, T is interpretable in S and vice-versa, i.e. if T and S are *mutually interpretable,* then one can conclude that T and S do not differ in strength. It is known that if T is interpretable in S and S is consistent then T must be consistent, too, and as already noted, if T is essentially undecidable then S must be essentially undecidable, too. The notion of interpretability, as well as the notion of essential undecidability and Robinson arithmetic itself, were defined in [TMR53]. For more information on the notion of interpretation see e.g. [Vis98]. As also already noted, Q^- is mutually interpretable with Q, see [Šve07].

3 An interpretation of Q^- in TC

In a series of lemmas, when saying that something is the case we mean "provably in TC", and by proofs we mean proofs in TC. Some of the statements in Lemma 1 also appeared in [GZ08].

Lemma 1 *(a)* $\forall x(x\mathtt{a} \neq \varepsilon \ \& \ \mathtt{a}x \neq \varepsilon)$. *The same is true for* $\mathtt{b}$ *and* $\mathtt{c}$.
(b) $\forall x \forall y(xy = \varepsilon \rightarrow x = \varepsilon \ \& \ y = \varepsilon)$.
(c) $\forall x \forall y(x\mathtt{a} = y\mathtt{a} \vee \mathtt{a}x = \mathtt{a}y \rightarrow x = y)$. *The same is true for* $\mathtt{b}$ *and* $\mathtt{c}$.
(d) $\forall x \forall y \forall u(u\mathtt{a} = xy \rightarrow y = \varepsilon \vee \exists y'(y = y'\mathtt{a}))$. *The same is true for* $\mathtt{b}$ *and* $\mathtt{c}$.

Proof (a) Assume $x\mathtt{a} = \varepsilon$. Then $\mathtt{b}x\mathtt{a} = \mathtt{b}$. By TC5, $\mathtt{b}x = \varepsilon$ or $\mathtt{a} = \varepsilon$. However, $\mathtt{a} = \varepsilon$ is not the case by TC4, while $\mathtt{b}x = \varepsilon$ yields $\mathtt{a} = \mathtt{b}$ by TC1, a contradiction with TC7.

(b) If $xy = \varepsilon$ then $xy\mathtt{a} = \mathtt{a}$. By TC4, $x = \varepsilon$ or $y\mathtt{a} = \varepsilon$. The latter is excluded by (a). From $x = \varepsilon$ we have $y\mathtt{a} = \mathtt{a}$. Using TC4 again, we have $y = \varepsilon$ or $\mathtt{a} = \varepsilon$. So $y = \varepsilon$.

(c) Assume $x\mathtt{a} = y\mathtt{a}$. By TC3 there exists a w such that $xw = y$ and $w\mathtt{a} = \mathtt{a}$, or $yw = x$ and $w\mathtt{a} = \mathtt{a}$. In both cases, from $w\mathtt{a} = \mathtt{a}$ we have $w = \varepsilon$. So $x = y$.

(d) Assume $u\mathtt{a} = xy$. By TC3 we have a w such that $uw = x \ \& \ wy = \mathtt{a}$, or $xw = u \ \& \ w\mathtt{a} = y$. In the second case we can take $y' := w$. So consider the first case, $uw = x \ \& \ wy = \mathtt{a}$. By TC4 we have $w = \varepsilon$ or $y = \varepsilon$. If $y = \varepsilon$ then we are done, and if $w = \varepsilon$ then for $y' := \varepsilon$ we have $y'\mathtt{a} = y$. ∎

We write $x \sqsubseteq y$ as a shortcut for $\exists s \exists t(sxt = y)$. We read $x \sqsubseteq y$ as *the string x is a substring of the string y,* or x *occurs* in y, or x has *occurences* in y. We write $x \sqsubset\!\!\!\cdot \ y$ for $\exists t(xt = y)$, i.e. to say that x is an *initial segment* of y or that y *begins* by x. Similarly, we write $x \sqsupset\!\!\!\cdot \ y$ to say that x is an *end segment* of y, i.e. that y *ends* by x. Using this notation, we can rewrite Lemma 1(d) as follows: $\mathtt{a} \sqsupset\!\!\!\cdot \ xy \rightarrow y = \varepsilon \vee \mathtt{a} \sqsupset\!\!\!\cdot \ y$. We know that if $x \sqsubset\!\!\!\cdot \ y$ or $x \sqsupset\!\!\!\cdot \ y$ then $x \sqsubseteq y$. It is easy to use Lemma 1(b) to show that if $x \sqsubseteq \mathtt{a}$ then $x = \varepsilon$ or $x = \mathtt{a}$, and if $x \sqsubseteq \varepsilon$ then $x = \varepsilon$.

Lemma 2 $\mathtt{a} \sqsubseteq xy \rightarrow \mathtt{a} \sqsubseteq x \vee \mathtt{a} \sqsubseteq y$. *The same is true for* $\mathtt{b}$ *and* $\mathtt{c}$.

Proof We have s and t such that $(s\mathtt{a})t = xy$. By TC3 there is a w such that $s\mathtt{a}w = x \ \& \ wy = t$, or $xw = s\mathtt{a} \ \& \ wt = y$. In the first case $\mathtt{a} \sqsubseteq x$. In the second case, from $xw = s\mathtt{a}$ and Lemma 1(d) we have $w = \varepsilon$ or $\mathtt{a} \sqsupset\!\!\!\cdot \ w$. If $w = \varepsilon$ then $xw = s\mathtt{a}$ yields $\mathtt{a} \sqsubseteq x$. If $\mathtt{a} \sqsupset\!\!\!\cdot \ w$ then $wt = y$ yields $\mathtt{a} \sqsubseteq y$. ∎

We say that x is a *number* and write $\mathrm{Num}(x)$ if each non-empty substring of x ends by $\mathtt{a}$. In symbols, $\mathrm{Num}(x) \equiv \forall u(u \sqsubseteq x \ \& \ u \neq \varepsilon \rightarrow \mathtt{a} \sqsupset\!\!\!\cdot \ u)$.

Lemma 3 *(a) Any substring of a number is a number. A number has no occurences of* $\mathtt{b}$ *or* $\mathtt{c}$.
(b) The strings ε *and* $\mathtt{a}$ *are numbers.*

(c) If x is a number and $x \neq \varepsilon$ then $x = y\mathtt{a}$ for some (number) y.
(d) If x and y are numbers then xy is a number.

Proof Verification of (a)–(c) is left to the reader. In (d), assume that x and y are numbers and u is a non-empty substring of xy. We have $sut = xy$ for some s and t. By axiom TC3, there is a w satisfying $suw = x \ \& \ wy = t$, or $xw = su \ \& \ wt = y$. In the first case u is a non-empty substring of x and thus must end by $\mathtt{a}$. In the second case, where $xw = su \ \& \ wt = y$, distinguish cases $w = \varepsilon$ and $w \neq \varepsilon$. If $w = \varepsilon$ then again, $u \sqsubseteq x$ and so $\mathtt{a} \sqsupset u$. If $w \neq \varepsilon$ then w is a non-empty substring of y. So $\mathtt{a} \sqsupset w$, i.e. $w = w'\mathtt{a}$ for some w'. Now from $xw'\mathtt{a} = su$ we have $\mathtt{a} \sqsupset u$ by Lemma 1(d). ∎

We take the formula $\mathrm{Num}(x)$ as the domain of the interpretation we construct, an interpretation of Q^- in TC. The domain is non-empty by Lemma 3(b). We define 0 as ε and, for a number x, $\mathrm{S}(x)$ as $x\mathtt{a}$. And we define the sum of numbers x and y to be the concatenation xy, i.e. we interpret $\mathrm{A}(x, y, z)$ as $xy = z$. By Lemma 3 (b) and (d), our domain is closed under both functions in the language of Q^-, i.e. 0 and S. Validity of axioms Q1-Q3 follows from Lemma 1(c), Lemma 1(a), and Lemma 3(c) respectively. Validity of axioms A, G4 and G5 is immediate. Note that, for the purpose of interpreting Q^- in TC, addition could have been a non-total function, but in our setting it *is total.* It remains to interpret multiplication.

Lemma 4 *(a) Assume that $s\mathtt{b}u = q\mathtt{b}x$ and u and x have no occurences of $\mathtt{b}$. Then $s = q$ and $u = x$. The same is true for $\mathtt{a}$ and $\mathtt{c}$.*
(b) Assume that $s\mathtt{b}u\mathtt{b}t = q\mathtt{b}x\mathtt{b}$ and u and x have no occurences of $\mathtt{b}$. Then either $s = q$, $u = x$, and $t = \varepsilon$, or there exists a w such that $s\mathtt{b}u\mathtt{b}w = q\mathtt{b}$ and $wx\mathtt{b} = t$.

Proof Apply axiom TC3 to $(s\mathtt{b})u = (q\mathtt{b})x$. There is a w such that $s\mathtt{b}w = q\mathtt{b}$ and $wx = u$, or $q\mathtt{b}w = s\mathtt{b}$ and $wu = x$. Consider the case $s\mathtt{b}w = q\mathtt{b}$ and $wx = u$, and note that the other case is symmetric. If $w \neq \varepsilon$ then $\mathtt{b} \sqsupset w$ by Lemma 1(d). Then from $wx = u$ we have $\mathtt{b} \sqsubseteq u$, a contradiction with the assumption that u has no occurences of $\mathtt{b}$. So $w = \varepsilon$. Then $x = u$, and from $s\mathtt{b} = q\mathtt{b}$ we have $s = q$ using Lemma 1(c).

In (b), apply axiom TC3 to $(s\mathtt{b}u\mathtt{b})t = (q\mathtt{b})(x\mathtt{b})$. If there is a w satisfying $s\mathtt{b}u\mathtt{b}w = q\mathtt{b}$ and $wx\mathtt{b} = t$ then we are done.

So suppose that we have a w such that $q\mathtt{b}w = s\mathtt{b}u\mathtt{b}$ and $wt = x\mathtt{b}$. We may assume that $w \neq \varepsilon$ since otherwise $s\mathtt{b}u\mathtt{b}\varepsilon = q\mathtt{b}$ and $\varepsilon x\mathtt{b} = t$, and we are done again. Then t must be empty: if $t \neq \varepsilon$ then from $wt = x\mathtt{b}$ and Lemma 1(d) we would have $t = t'\mathtt{b}$ for some t', from $w \neq \varepsilon$, Lemma 1(d), and $q\mathtt{b}w = s\mathtt{b}u\mathtt{b}$ we have $w = w'\mathtt{b}$ for some w', and then $w'\mathtt{b}t'\mathtt{b} = x\mathtt{b}$ and Lemma 1(c) would yield $w'\mathtt{b}t' = x$, a contradiction with the assumption that x has no occurences of $\mathtt{b}$.

From $t = \varepsilon$ we have $w = x\mathtt{b}$ and $q\mathtt{b}x\mathtt{b} = s\mathtt{b}u\mathtt{b}$. Then, using Lemma 1(c) and using (a) already proved, $q = s$ and $x = u$ follows. ∎

We say that w is a *(product) witness* for $x \times y$ and write $\mathrm{PWitn}(x, y, w)$ if the following conditions are true:

(i) the strings x and y are numbers,
(ii) there is a number z such that $\mathtt{b}y\mathtt{c}z\mathtt{b} \sqsupset w$,
(iii) $\forall u_2 \forall v_2 \forall s \forall t (s\mathtt{b}u_2\mathtt{c}v_2\mathtt{b}t = w \ \& \ \mathrm{Num}(u_2) \ \& \ \mathrm{Num}(v_2) \ \& \ s \neq \varepsilon \rightarrow \exists u_1 \exists v_1 (\mathrm{Num}(u_1) \ \& \ \mathrm{Num}(v_1) \ \& \ u_2 = u_1\mathtt{a} \ \& \ v_2 = v_1 x \ \& \ \mathtt{b}u_1\mathtt{c}v_1 \sqsupset s))$,
(iv) $\mathtt{bcb} \sqsubset w$.

The formula $\mathrm{PWitn}(x, y, w)$ roughly says that "w ends by $\mathtt{b}y\mathtt{c}z\mathtt{b}$, begins by $\mathtt{bcb}$, and each its substring $\mathtt{b}u_2\mathtt{c}v_2\mathtt{b}$, which is not an initial segment of w, is immediately preceded by $\mathtt{b}u_1\mathtt{c}v_1$, where $u_2 = u_1 + 1$ and $v_2 = v_1 + x$." So for example, $\mathtt{bcbacaabaacaaaabaaacaaaaaab}$ is a product witness for 2×3.

Lemma 5 *Let x and y be numbers.*
(a) $\mathrm{PWitn}(x, \varepsilon, \mathtt{bcb})$.
(b) $\forall w' (\mathrm{PWitn}(x, \varepsilon, w') \rightarrow w' = \mathtt{bcb})$.
(c) $\forall q \forall z (\mathrm{PWitn}(x, y, q\mathtt{b}y\mathtt{c}z\mathtt{b}) \rightarrow \mathrm{PWitn}(x, y\mathtt{a}, q\mathtt{b}y\mathtt{c}z\mathtt{b}y\mathtt{ac}zx\mathtt{b}))$.
(d) Let w' be a witness for $x \times y\mathtt{a}$. Then there is a string q' and a number v such that $w' = q'\mathtt{b}y\mathtt{c}v\mathtt{b}y\mathtt{ac}vx\mathtt{b}$, where $\mathrm{PWitn}(x, y, q'\mathtt{b}y\mathtt{c}v\mathtt{b})$.

Proof In (a), where in addition $y = \varepsilon$, the string $\mathtt{bcb}$ evidently ends by a string $\mathtt{b}y\mathtt{c}z\mathtt{b}$ where z is a number, and begins by $\mathtt{bcb}$. So, in the definition of product witness, it remains to verify the condition (iii). Let u_2, v_2, s, t be such that u_2 and v_2 are numbers and $s\mathtt{b}u_2\mathtt{c}v_2\mathtt{b}t = \mathtt{bcb}$. Repeated use of Lemma 1 (d) and (c), axiom TC5 and Lemma 1(a) shows that t, v_2, u_2, and s must all be empty. So $\mathtt{bcb}$ cannot be written as $s\mathtt{b}u_2\mathtt{c}v_2\mathtt{b}t$ with non-empty s, and thus condition (iii) is satisfied.

We omit the proof of (b) as similar to the proof of (d) given below. In (c), assume that $q\mathtt{b}y\mathtt{c}z\mathtt{b}$ is a product witness for $x \times y$. Think about the string $q\mathtt{b}y\mathtt{c}z\mathtt{b}y\mathtt{ac}zx\mathtt{b}$. The strings $y\mathtt{a}$ and zx are numbers by Lemma 3(d); so conditions (i) and (ii) are satisfied w.r.t. x and $y\mathtt{a}$. Also (iv) is satisfied because already $q\mathtt{b}y\mathtt{c}z\mathtt{b}$ begins by $\mathtt{bcb}$. It remains to verify the condition (iii). So consider $s \neq \varepsilon$ and t and numbers u_2, v_2 such that $s\mathtt{b}u_2\mathtt{c}v_2\mathtt{b}t = q\mathtt{b}y\mathtt{c}z\mathtt{b}y\mathtt{ac}zx\mathtt{b}$. By Lemma 2, the strings $u_2\mathtt{c}v_2$ and $y\mathtt{ac}zx$ have no occurences of $\mathtt{b}$. So Lemma 4(b) can be used as follows:

$$s\mathtt{b}\underbrace{u_2\mathtt{c}v_2}_{u}\mathtt{b}t = \underbrace{q\mathtt{b}y\mathtt{c}z}\,\mathtt{b}\,\underbrace{y\mathtt{ac}zx}\,\mathtt{b}.$$

In the first case, where $s = q\mathtt{b}y\mathtt{c}z$, $u_2\mathtt{c}v_2 = y\mathtt{ac}zx$, and $t = \varepsilon$, one can easily use Lemma 4(a) and conclude that $u_2 = y\mathtt{a}$ and $v_2 = zx$. So indeed, s ends by $\mathtt{b}u_1\mathtt{c}v_1$ where $u_2 = u_1\mathtt{a}$ and $v_2 = v_1 x$. In the second case we have a w such that $s\mathtt{b}u_2\mathtt{c}v_2\mathtt{b}w = q\mathtt{b}y\mathtt{c}z\mathtt{b}$ and $wy\mathtt{ac}zx\mathtt{b} = t$. By the assumption that $q\mathtt{b}y\mathtt{c}z\mathtt{b}$ is a product witness for $x \times y$, the string s must end by $\mathtt{b}u_1\mathtt{c}v_1$ as required.

Finally, to prove (d), assume that w' is a witness for $x \times y\mathtt{a}$. We know that w' ends by $\mathtt{b}y\mathtt{ac}v'\mathtt{b}$ where v' is a number, and begins by $\mathtt{bcb}$. So we have strings t and q'' and a situation where Lemma 4(b) can be used as follows:

$$w' = \underbrace{\quad}_{s}\mathtt{b}\underbrace{\mathtt{c}}_{u}\mathtt{b}t = q''\mathtt{b}\underbrace{y\mathtt{ac}v'}_{x}\mathtt{b}.$$

The case where $\varepsilon = q''$, $\mathtt{c} = y\mathtt{ac}v'$, $t = \varepsilon$ is impossible, $\mathtt{c}$ cannot have a substring $\mathtt{ac}$. So we have a w such that $\mathtt{bcb}w = q''\mathtt{b}$ and $wy\mathtt{ac}v'\mathtt{b} = t$. From $\mathtt{bcb}w = q''\mathtt{b}$ one can conclude $q'' \neq \varepsilon$. Since w' is a witness, condition (iii) says that $q'' = q'\mathtt{b}u_1\mathtt{c}v$ for some q' and some numbers u_1 and v such that $y\mathtt{a} = u_1\mathtt{a}$ and $v' = vx$. Then $y = u_1$ and $w' = q'\mathtt{b}y\mathtt{c}v\mathtt{b}y\mathtt{ac}vx\mathtt{b}$. Evidently, $q'\mathtt{b}y\mathtt{c}v\mathtt{b}$, which is the same as $q''\mathtt{b}$, satisfies all conditions (i)–(iv) in the definition of a witness for $x \times y$. ∎

Having Lemma 5, we can define the formula $\mathrm{M}(x, y, z)$, saying that z is a product of x and y, as follows:

$$\mathrm{M}(x,y,z) \equiv \exists w(\mathrm{PWitn}(x,y,w)\ \&\ \forall w'(\mathrm{PWitn}(x,y,w') \rightarrow w' = w)\ \& \\ \&\ \mathtt{c}z\mathtt{b} \sqsupset w).$$

Theorem *The theory* Q^- *is interpretable in* TC. *Thus also Robinson arithmetic* Q *is interpretable in* TC, *and* TC *is essentially undecidable.*

Proof It remains to consider axioms about multiplication, i.e. M, G6, and G7. If $\mathrm{M}(x, y, z_1)$ and $\mathrm{M}(x, y, z_2)$, then there is a w that is the unique witness for $x \times y$ and such that $\mathtt{b}y\mathtt{c}z_1\mathtt{b} \sqsupset w$ and $\mathtt{b}y\mathtt{c}z_2\mathtt{b} \sqsupset w$. Then the usual argument, i.e. Lemma 1(c) and Lemma 4(a), shows that $z_1 = z_2$. So validity of axiom M in our interpretation follows. Validity of axiom G6 follows from Lemma 5 (a) and (b). Consider axiom G7. Let $\mathrm{M}(x, y, z)$ and $\mathrm{A}(z, x, u)$. We have to verify $\mathrm{M}(x, \mathrm{S}(y), u)$. According to our definitions, $\mathrm{A}(z, x, u)$ says $zx = u$, while $\mathrm{S}(y)$ is $y\mathtt{a}$. We know from $\mathrm{M}(x, y, z)$ that there exists a unique witness for $x \times y$; it must have the form $q\mathtt{b}y\mathtt{c}z\mathtt{b}$. Then Lemma 5(c) says that $q\mathtt{b}y\mathtt{c}z\mathtt{b}y\mathtt{ac}zx\mathtt{b}$ is a witness for $x \times y\mathtt{a}$. To verify that it is the only witness, let w' be a witness for $x \times y\mathtt{a}$. By Lemma 5(d), $w' = q'\mathtt{b}y\mathtt{c}v\mathtt{b}y\mathtt{ac}vx\mathtt{b}$ where v is a number and $q'\mathtt{b}y\mathtt{c}v\mathtt{b}$ is a witness for $x \times y$. However, we know that $q\mathtt{b}y\mathtt{c}z\mathtt{b}$ is the only witness for $x \times y$. Thus $q'\mathtt{b}y\mathtt{c}v\mathtt{b} = q\mathtt{b}y\mathtt{c}z\mathtt{b}$. Then $q'\mathtt{b}y\mathtt{c}v = q\mathtt{b}y\mathtt{c}z$, and Lemma 4(a) says $v = z$ and $q' = q$. Thus indeed, $w' = q\mathtt{b}y\mathtt{c}z\mathtt{b}y\mathtt{ac}zx\mathtt{b}$. ∎

Petr Hájek considered a somewhat stronger variant Q^h of Q^-, having the same language and similar axioms, but with equivalences instead of implications in axioms G5 and G7, see [Háj07]. So in Q^h, if $\mathrm{S}(y)$ can be added to x from the right then also y can be added to x from the right, and if x can be multiplied by $\mathrm{S}(y)$ from the right then x can also be multiplied by y, and their product can be added to x from the left. One can verify that Hájek's axioms are valid in our interpretation as well. Albert Visser noticed that there exists a simple

interpretation of Q in Q^h, one that does not use the Solovay's technique of shortening of cuts: it is basically sufficient to introduce an "ideal" individual ∞ and stipulate that ∞ is the new sum or product of x and y whenever the old sum or product of x and y do not exist. So since the Solovay's technique is an essential ingredient of [Šve07], "Visser's detour" via Q^h yields a more straightforward interpretation of Q in TC than the detour via Grzegorczyk's Q^-.

Note that TC is easily interpretable in the bounded arithmetic $\mathrm{I}\Delta_0$. Since $\mathrm{I}\Delta_0$ is known to be interpretable in Q, all theories TC, Q^-, Q^h, Q, and $\mathrm{I}\Delta_0$ are mutually interpretable.

References

[ČPR+07] V. Čačić, P. Pudlák, G. Restall, A. Urquhart, and A. Visser. Decorated linear order types and the theory of concatenation. Logic Group Preprint Series 258, Utrecht University, Utrecht, 2007.

[Gan09] M. Ganea. Arithmetic on semigroups. *J. Symb. Logic*, 74(1):265–278, 2009.

[Grz05] A. Grzegorczyk. Undecidability without arithmetization. *Studia Logica*, 79(2):163–230, 2005.

[GZ08] A. Grzegorczyk and K. Zdanowski. Undecidability and concatenation. In A. Ehrenfeucht, V. W. Marek, and M. Srebrny, editors, *Andrzej Mostowski and Foundational Studies*, pages 72–91. IOS Press, Amsterdam, 2008.

[Háj07] P. Hájek. Mathematical fuzzy logic and natural numbers. *Fundamenta Informaticae*, 81(1–3):155–163, 2007.

[Qui46] W. V. O. Quine. Concatenation as a basis for arithmetic. *J. Symb. Logic*, 11(4):105–114, 1946.

[Šve07] V. Švejdar. An interpretation of Robinson arithmetic in its Grzegorczyk's weaker variant. *Fundamenta Informaticae*, 81(1–3):347–354, 2007.

[TMR53] A. Tarski, A. Mostowski, and R. M. Robinson. *Undecidable Theories*. North-Holland, Amsterdam, 1953.

[Vis98] A. Visser. An overview of interpretability logic. In M. Kracht, M. Rijke, H. Wansing, and M. Zakharyaschev, editors, *Advances in Modal Logic 1 (AiML'96)*, volume 87 of *CSLI Lecture Notes*, pages 307–359, Berlin, October 1998. CSLI Publ., 1998.

[Vis09] A. Visser. Growing commas: a study of sequentiality and concatenation. *Notre Dame J. Formal Logic*, 50(1):61–85, 2009.